Using Geometric Brownian Motion to Simulate Stock Prices: Parameter Estimation and Model Validation

Samuel Jones

August 2025

Introduction

Geometric Brownian Motion (GBM) is one of the most widely used models in finance for simulating stock prices. It forms the basis of the famous Black-Scholes model and is popular because it's relatively simple to understand and implement. This report investigates the implementation and validation of GBM models for stock price simulation.

The main goals of the project are to: build a working GBM simulation using real stock data using rolling parameter estimation, evaluate model performace using visual and numeric metrics and to examine the trade offs between model accuracy and overfitting.

1 Theory Behind Geometric Brownian Motion

Geometric Brownian Motion is a mathematical model that tries to capture how stock prices move over time. It assumes that prices grow exponentially on average (the "drift") but with random fluctuations added in (the "volatility"), following the formula

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t, \tag{1}$$

where:

- S_t is the stock price at time t
- μ is the drift coefficient (expected return)
- σ is the volatility (standard deviation of returns)
- W_t is a standard Brownian motion (Wiener process)

1.1 Discrete-Time Implementation

For practical purposes, GBM is implemented using its discrete-time approximation. Given a time step Δt, the price evolution is given by:

$$S_{t+\Delta t} = S_t \cdot \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t} \cdot Z_t\right], \tag{2}$$

where $Z_t \sim \mathcal{N}(0,1)$ are independent standard normal random variables. This ensures that prices remain positive and follow a log-normal distribution.

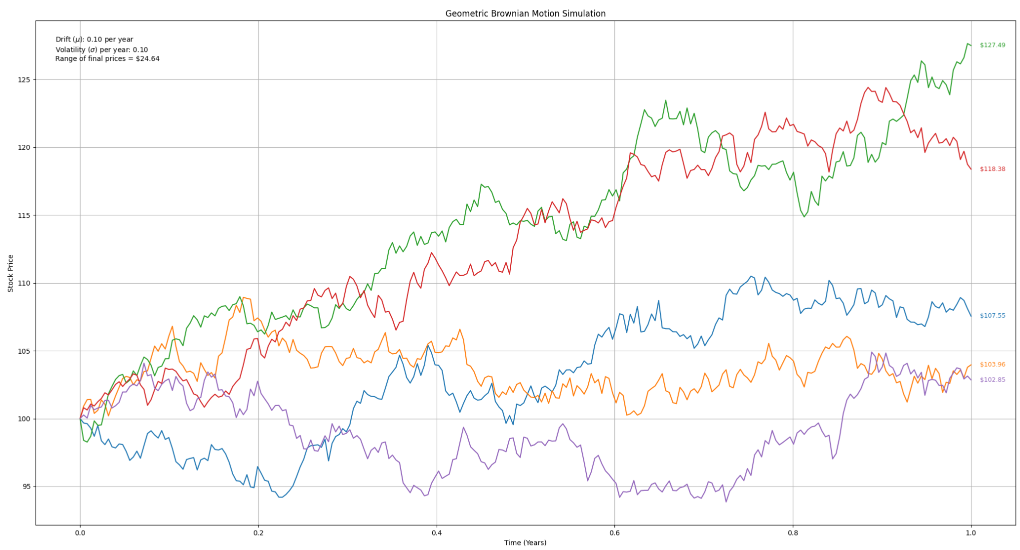

Figure 1: Five different GBM simulations over a 10-year period with drift $\mu = 0.1$ and volatility $\sigma = 0.1$.

Figure 1 shows what GBM looks like with fixed parameters, where the simulations exhibit realistic stock-like behaviour with clear upward trends and random path variations. Notably, longer time horizons with fixed parameters result in wider ranges of terminal values, and volatility tends to dominate drift effects over short to medium time scales.

1.2 GBM Limitatons

While GBM provides mathematical tractability and intuitive interpretation, it suffers from several notable limitations. The model assumes constant drift (μ) and volatility (σ) parameters, which contradicts the time-varying nature of real markets. Additionally, GBM presumes normally distributed returns and log-normal price distributions, failing to capture the heavy tails and skewness commonly observed in empirical data. The model also cannot account for sudden price jumps or discontinuities, as it generates only continuous paths. These limitations can lead to systematic underestimation of risk, particularly during extreme market conditions, motivating the development of more sophisticated models such as stochastic volatility or jump-diffusion processes.

2 Building the Model

2.1 Data and Parameter Estimation

This project used historical data from the SPDR S&P 500 ETF (SPY) downloaded using the yfinance library. Rather than using fixed parameters, a rolling window approach was used on historical log returns, with the model assuming 252 trading days per year.

For each time step t, the drift and volatility parameters are estimated using a rolling window of size w:

$$\hat{\mu}_t = 252 \cdot \frac{1}{w} \sum_{i=t-w}^{t-1} r_i \tag{3}$$

$$\hat{\sigma}_t = \sqrt{252} \cdot \sqrt{\frac{1}{w-1} \sum_{i=t-w}^{t-1} (r_i - \bar{r})^2} \tag{4}$$

where $r_i = \ln(S_i / S_{i-1})$ represents the log return at time i.

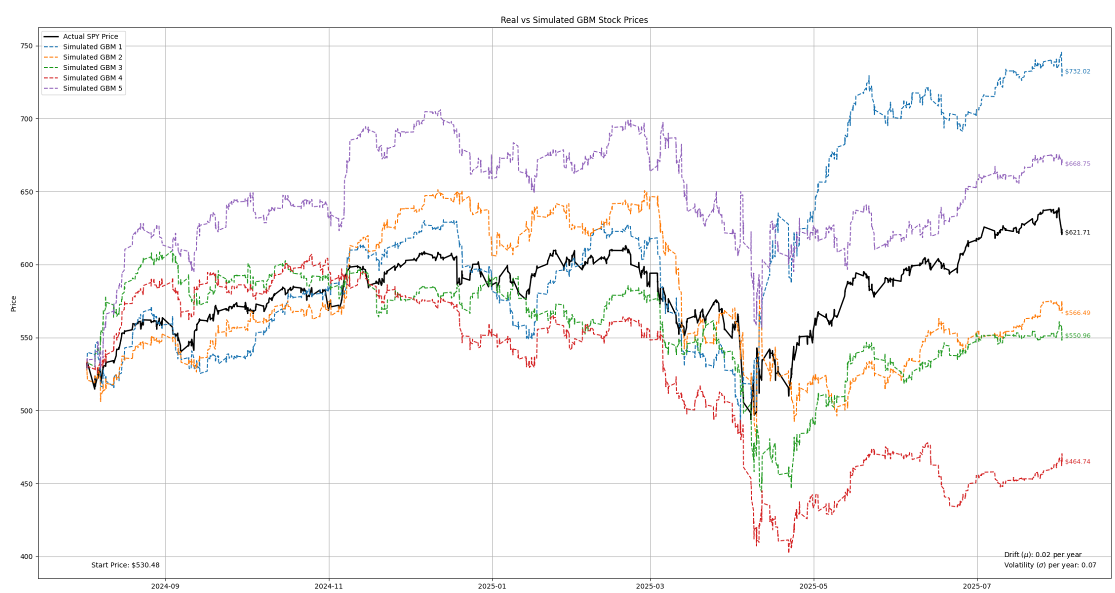

Figure 2: Five individual GBM simulations compared to the actual SPY price over one year, using a 21-day rolling window. The simulations roughly follow the real price but show quite a bit of variation.

Figure 2 demonstrates the behaviour of individual GBM simulations using a 21-day rolling parameter window. While the simulations generally follow the actual stock price, individual paths show significant range between them, with an end price range of $267.28. This inherent randomness of GBM requires the use of averaging over multiple models for a more meaningful trend analysis. This is shown in Figure 3 where the mean of 10,000 GBM simulations is compared with real SPY prices.

Figure 3: Average of 10,000 GBM simulations using a 21 day rolling window compared to actual SPY prices. Taking the average gives a much smoother result that follows the real price more closely.

3 Model Performance

3.1 Root Mean Squared Error Analysis

To quantitatively assess model performance, the Root Mean Squared Error (RMSE) was calculated between the mean and actual stock prices across different rolling window sizes. Table 1 shows these results.

Window Size (Days)	**RMSE ($)**
42	13.85
21	9.17
12	7.47
4	5.92
2	2.78

Table 1: RMSE values for different rolling window sizes using 10,000 simulations with standardized random seeds over identical time periods.

The results show a clear inverse relationship between window size and RMSE, with shorter windows producing apparently superior fits. This improvement however, comes at the cost of model stability and economic realism.

3.2 The Overfitting Problem

While minimizing RMSE represents a natural optimisation objective, the pursuit of low RMSE in financial modelling can lead to overfitting and unrealistic parameter estimates. Figure 4 demonstrates this using a 2-day rolling window.

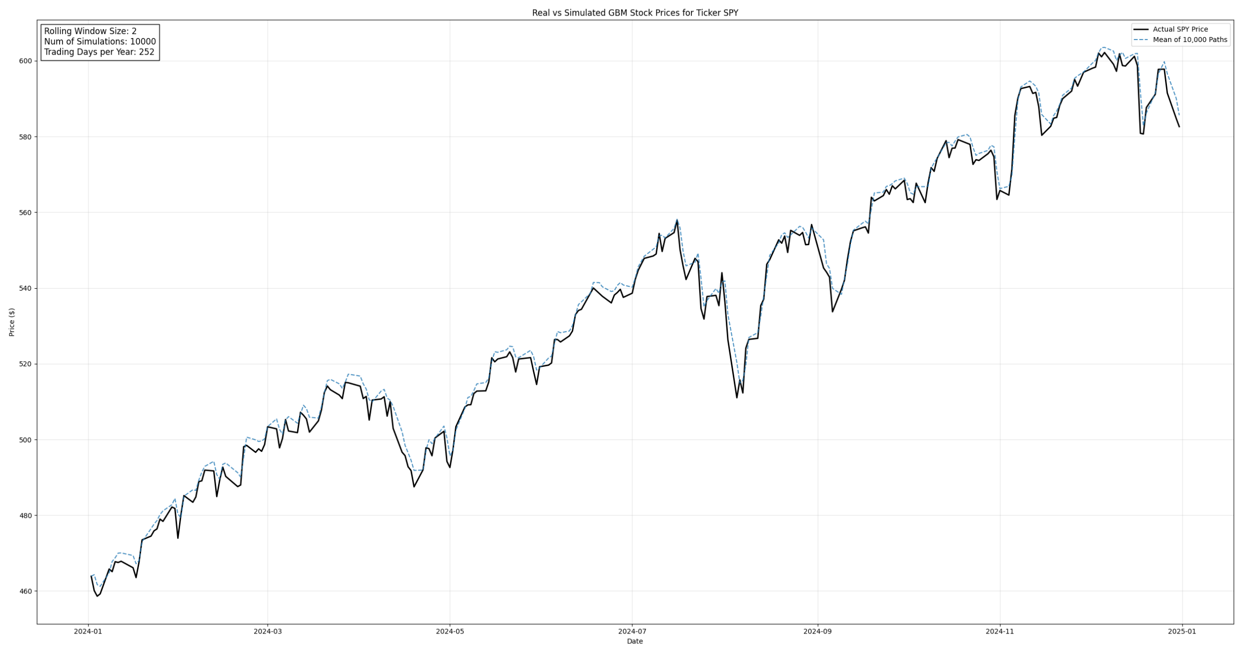

Figure 4: Average of 10,000 GBM simulations using a 2-day rolling window. Despite having the lowest RMSE ($2.78), this model is overfitted.

The 2-day window produces economically implausible parameter estimates, with annual drift rates ranging from -600% to +420%. This extreme parameter sensitivity reflects overfitting to recent price movements rather than genuine predictive capability. The apparently superior RMSE performance masks fundamental model instability that would prove catastrophic in real-world applications.

The inherent unpredictability of financial markets suggests that empirical trend capture and proper uncertainty quantification provide more valuable performance metrics than pure RMSE minimization. Models achieving very low RMSE may be overfitted or making unrealistic precision claims, whereas models accurately capturing market dynamics and volatility patterns—despite higher point prediction errors—offer greater practical value for risk management and investment decisions.

3.3 Using a Naïve RMSE

Since minimising RMSE on its own is not a good metric of model performance, a comparison can be introduced, the naïve RMSE, so called, as it assumes that the price at t + 1 is equal to the price at t, effectively predicting no change. This gives a benchmark to compare the model's RMSE to, if it is lower than the naïve RMSE, the model adds value, if it is not much better, or worse, the model does not add value.

Applied to SPY data from 2024-01-01 to 2025-01-01, the naïve RMSE was $4.27, while the GBM model yielded an RMSE of $8.33. While this is just one result, from testing using different time frames, the RMSE was typically aroud x2-3 larger than the naïve RMSE. This result indicates that the current GBM implementation may be too simplistic of a model, as it is unable to outperform the naïve baseline.

4 Confidence Intervals and Empirical Coverage

4.1 Building Confidence Intervals

Confidence Intervals (CI) provide a statistical framework for quantifying simulation uncertainty. For a given confidence level (e.g., 90%), intervals are constructed by computing the appropriate percentiles of simulated prices at each time step. A 90% confidence interval uses the 5th and 95th percentiles, representing the range within which 90% of simulated outcomes fall under the model assumptions.

4.2 Empirical Coverage

Empirical Coverage (EC) measures the proportion of time steps for which actual stock prices fall within the predicted confidence intervals. This metric serves as a diagnostic tool for model calibration, a well-calibrated models should exhibit empirical coverage closely matching the confidence level.

Figure 5 shows that for the GBM models used, a 90% CI results in a 100% EC, suggesting an overestimation of model volatility. To achieve a ∼90% EC, the confidence level had to be reduced to 35%, implying that the model has a far greater confidence that it's predictions statistically should.

4.3 Rectifying Volatility Over Estimates

A straightforward method to reduce the volatility input in the simulation is to apply a scaling factor directly to the estimated values. However this adjustment is not derived from the underlying data but introduced externally as a model tuning parameter. For instance, applying a volatility scaling factor of 0.3 resulted in an empirical coverage of 90.1% for a 90% confidence interval.

Replacing the simple rolling window with an exponential weighted moving average can rectify the volatility over estimates as they favour more recent data, not allowing large volatility spikes from early on in the rolling window to have such a large effect on volatility estimates.

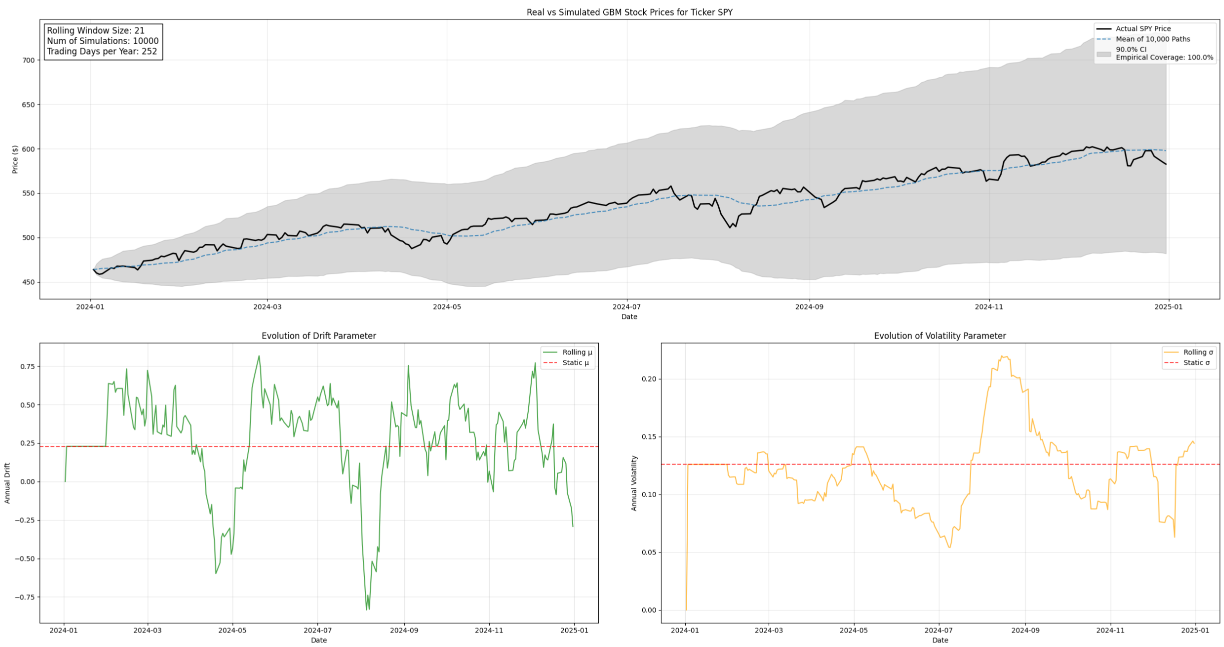

Figure 5: GBM simulation with 90% confidence intervals showing 100% empirical coverage, indicating overestimation of model uncertainty.

5 Conclusion

This investigation of Geometric Brownian Motion for stock price simulation reveals both the model's utility and its fundamental limitations when applied to real market data. While GBM provides a mathematical framework with intuitive interpretation, its practical implementation exposes several critical challenges that constrain its effectiveness as a predictive tool.

The rolling parameter estimation approach successfully addresses GBM's assumption of constant parameters, allowing the model to adapt to changing market conditions. However, this adaptation comes with a crucial trade-off between model responsiveness and stability. Shorter rolling windows (2-4 days) produce lower RMSE values but generate economically implausible parameter estimates, with annual drift rates reaching extremes of -600% to +420%. Conversely, longer windows (21-42 days) provide more stable parameters but at the cost of reduced adaptability to recent market movements.

The comparison against naïve RMSE benchmarks reveals that the GBM model consistently underperforms the simple assumption that prices remain unchanged, with RMSE values typically 2-3 times larger than the naïve baseline. This result suggests that the current implementation, despite its mathematical sophistication, may be too simplistic to capture the complex dynamics of real financial markets.

The use of confidence intervals and empirical coverage further highlights the model's calibration issues. The systematic overestimation of volatility, evidenced by 100% empirical coverage for comparatively small confidence intervals, indicates that the model is overestimating volatility. While scaling factors can mechanically correct this issue, such adjustments represent external model tuning rather than improvements taken from market data.

GBM's limitations—constant parameters, normal return distributions, and inability to capture price jumps—become particularly apparent when compared with real market data., with the model's systematic underperformance relative to naïve benchmarks suggesting that financial markets exhibit complexities that simple diffusion processes cannot adequately capture.

Future research directions should focus on addressing these limitations through more sophisticated approaches. Stochastic volatility models such as the Heston model, which allows volatility itself to follow a mean-reverting stochastic process, or the Black-Scholes-Merton (BSM) jump-diffusion framework, which incorporates sudden price discontinuities, may provide better empirical fit while maintaining interpretability. Regime-switching models could also capture the time-varying nature of market conditions that GBM's constant parameters cannot accommodate. Additionally, exploring alternative parameter estimation techniques, such as exponentially weighted moving averages or Bayesian updating methods, could improve model calibration without sacrificing stability.

While GBM remains valuable for theoretical analysis and educational purposes, practitioners should exercise caution when using it for real-world applications. The model's primary value may lie not in precise price prediction but in providing a baseline framework for understanding market dynamics and quantifying uncertainty—albeit with the recognition that actual market behavior is likely to be more complex than any simple mathematical model can fully capture.