

Black-Scholes Extensions in Deep Learning Option Pricing

Zongrun Li

School of Risk and Actuarial Studies, UNSW Business School, Sydney, Australia
z5638221@ad.unsw.edu.au

Abstract. Option pricing is important in modern financial markets because options are widely used for hedging, speculation, and structured investment. The Black-Scholes model remains the key benchmark, but its assumptions of constant volatility, frictionless trading, and lognormal returns often differ from observed market behaviour. This review examines how recent deep learning studies extend Black-Scholes-based pricing and hedging. Using a comparative literature approach, it first discusses Black-Scholes as a theoretical benchmark, then analyses hybrid deep learning models for pricing accuracy, and finally evaluates residual-learning methods for delta hedging. The review finds that Black-Scholes is best treated as an interpretable formula-based benchmark rather than an empirical prediction model. Because the original model is a mathematical pricing formula, it does not by itself provide MSE or Gain Ratio; those metrics are introduced by later empirical studies that compare against it. Recent evidence shows that hybrid neural networks can reduce pricing MSE by 94.5% relative to a Black-Scholes benchmark based on historical volatility, while residual learning can improve hedging performance by more than 100% relative to direct learning when the benchmark uses implied Black-Scholes delta. The paper concludes that future work should combine financial structure with data-driven correction rather than replace Black-Scholes with a pure black-box model.

Keywords: option pricing, Black-Scholes model, deep learning, delta hedging, mean squared error

1. Introduction

Options are among the most important derivative instruments in modern financial markets. They allow investors, firms, and financial institutions to transfer risk, protect portfolios, build leveraged positions, and design payoff structures that cannot be achieved by simply buying or selling the underlying asset. Because of this, option pricing is not only a technical valuation problem. It is also closely connected to market efficiency, risk control, trading strategy, and financial engineering. A small pricing error can influence hedging decisions, margin requirements, risk exposure, and portfolio performance. For this reason, researchers and practitioners have continued to study option pricing for more than fifty years.

The foundation of modern option pricing is the Black-Scholes model [1]. Its importance goes beyond the well-known closed-form formula for European call options. More fundamentally, it introduced a no-arbitrage pricing framework in which an option can be valued through a riskless

hedged portfolio consisting of the option and the underlying asset. Merton's continuous-time pricing theory extended this framework and strengthened its theoretical foundation [2]. Later, the binomial tree model provided a more intuitive discrete-time version of the same no-arbitrage idea [3]. Together, these classical models created the basis for later studies on stochastic volatility, implied volatility, volatility smiles, and dynamic hedging.

However, real-world option markets are much more complicated than the assumptions used in classical theory. The Black-Scholes model assumes constant volatility, continuous trading, no transaction costs, a constant risk-free interest rate, and lognormally distributed asset prices. In practice, these assumptions are often violated. Implied volatility changes across strike prices and maturities, creating volatility smiles and skews. This means that a single constant-volatility parameter cannot fully describe market expectations. Stochastic-volatility models were introduced to relax part of this limitation, while early neural-network studies showed that non-parametric learning methods could capture option-pricing relationships directly from market data [4, 5]. More recently, deep learning has been applied not only to option pricing but also to hedging, risk control, and dynamic portfolio adjustment [6-11].

This review addresses a narrower gap: Black-Scholes and deep learning are often presented as substitutes, although recent evidence suggests that they are more useful when combined. Using a comparative literature review, this paper examines three stages: Black-Scholes as a benchmark, hybrid deep learning for pricing, and residual learning for delta hedging. It compares their contributions using metrics appropriate to each setting, including pricing MSE, mean squared hedging error, and Gain Ratio.

2. Black-Scholes as the benchmark model

The Black-Scholes model should be understood as the core benchmark of modern option-pricing theory. Under its assumptions, the price of a European option is derived by constructing a continuously rebalanced portfolio of the option and the underlying asset. If this portfolio is riskless over a very small time interval, it must earn the risk-free rate. Otherwise, an arbitrage opportunity would exist. This reasoning makes option pricing independent of investors' individual risk preferences. As a result, option valuation is transformed from a subjective forecasting problem into a no-arbitrage consistency condition [1].

The model uses only a limited number of inputs: the current underlying asset price, the strike price, the time to expiration, the risk-free interest rate, and volatility. This simple structure explains why the model has remained widely used in both academic research and financial practice. It allows users to calculate a theoretical option value quickly. It also provides important risk measures, especially delta, which measures the sensitivity of the option price to changes in the underlying asset price. The same logic can also be applied outside standard option markets. For example, in corporate finance, equity can be interpreted as a call option on the value of a firm's assets [1, 2].

For this review, it is important to state this precisely: the original Black-Scholes framework is a mathematical pricing model, not an empirical machine-learning algorithm or software package. It therefore does not itself report MSE, RMSE, accuracy, or Gain Ratio. Those indicators only appear when later researchers implement the formula as a benchmark and compare its pricing or hedging output with market data. Its main contribution is the theoretical benchmark created by no-arbitrage pricing, which later empirical models can use when evaluating pricing errors, hedging errors, or model improvements.

At the same time, the model's limitations are well known. Volatility is assumed to be constant, but market prices imply different volatilities for different strike prices and maturities. Trading is

assumed to be continuous and costless, while real hedging is discrete and involves bid-ask spreads, transaction costs, liquidity constraints, and market impact. Asset returns may also show jumps, heavy tails, and stochastic volatility. These problems do not make Black-Scholes useless. Instead, they explain why many later models keep Black-Scholes as a baseline and then add extra layers through stochastic volatility, machine learning, or deep hedging methods.

3. Deep learning extensions for pricing accuracy

One major recent research direction is to improve Black-Scholes pricing accuracy through deep learning. In this setting, the objective changes from deriving a theoretical no-arbitrage price to predicting market prices more accurately. The predicted option price is compared with the observed market price, and the pricing difference is evaluated using statistical indicators such as mean squared error. This is different from the classical Black-Scholes approach, which focuses on theoretical consistency under simplified assumptions.

A representative hybrid study compares several deep learning architectures with a Black-Scholes benchmark using S&P 500 ETF options, specifically SPY options [8]. In that study, the observed option price is calculated as the average of bid and ask prices. Options with zero trading volume are excluded because inactive quotes may not reflect reliable market prices. The Black-Scholes benchmark is calculated using volatility estimated from annualized 20-day intraday returns. The neural-network inputs include option type, underlying price, time to expiration, volatility, option volume, the strike-to-underlying-price ratio, the risk-free rate, and the Black-Scholes calculated price.

The numerical results are important because they show whether deep learning produces a measurable improvement over a closed-form benchmark. The standalone Black-Scholes benchmark reports an average MSE of 10.9518. A single-stage deep multilayer perceptron reduces the average MSE to 5.6699. The best-performing optimized classifier-pricer model further reduces the MSE to 0.59810. This equals a 94.5% reduction relative to the Black-Scholes benchmark [8]. In practical terms, the deep learning model not only appears more flexible; it also produces a much smaller squared pricing error.

This improvement is possible because deep learning can capture nonlinear relationships among pricing variables. The classifier-pricer design is also useful because option contracts do not behave identically across all regions. Near-the-money, deep in-the-money, and out-of-the-money options may differ in liquidity, volatility-smile patterns, and pricing errors. In the optimized classifier-pricer model, contracts are first separated by Black-Scholes error thresholds into four classes, and a separate DMLP pricer is then trained for each class. A modular model can therefore specialize across option groups rather than forcing a single formula to fit all contracts in the same way.

However, this approach also has limitations. Deep learning models are usually less transparent than closed-form pricing formulas. Their performance depends heavily on data quality, sample selection, feature construction, and model training. Results based on SPY options and closing-time pricing conditions may not generalize perfectly to all assets, markets, or intraday trading environments. Therefore, deep learning should not be treated as a simple replacement for financial theory. It is more useful when it is combined with Black-Scholes information and evaluated carefully across different market conditions.

4. Deep learning extensions for delta hedging

A second important research direction shifts the focus from pricing to hedging. In practice, an option trader may care not only about whether a model matches the observed market price, but also about whether the hedge created by the model reduces profit-and-loss risk. The Black-Scholes delta remains the standard benchmark for this task because it measures how sensitive an option price is to movements in the underlying asset. However, when the assumptions of Black-Scholes fail, a standard delta hedge can still generate significant hedging errors.

A recent residual-learning approach keeps the implied Black-Scholes delta as the economic baseline and trains a feedforward neural network to learn only the residual correction [11]. This is an effective compromise between theory and data. Instead of forcing the neural network to learn the entire hedging function from the beginning, the model starts with the Black-Scholes delta and then learns the difference between that delta and the empirically better hedge ratio. The authors argue that the residual function is smoother than the full hedging function, especially close to expiration, where option payoffs and derivatives can become discontinuous.

The empirical design is strong because it uses ten years of daily S&P 500 index option data from 2010 to 2019. The dataset includes bid and ask quotes, implied volatility, and option Greeks such as delta, theta, vega, and gamma. After filtering untraded contracts, missing observations, very short maturities, and extreme delta values, the final dataset contains more than 2.073 million quotations [11]. This large sample gives the hedging evidence more credibility than a small experimental dataset.

The main evaluation metric is the mean squared one-step local hedging error. This measures the squared difference between the option value change and the hedge position multiplied by the change in the underlying asset price. The improvement over the Black-Scholes delta benchmark is reported through Gain Ratio, which measures how much the proposed model reduces mean squared hedging error relative to the benchmark. The results show that residual learning often improves hedging performance by more than 100% compared with directly learning the full hedging function. When time to maturity and Black-Scholes delta are included in the feature vector, the residual model outperforms the corresponding direct-learning model by 129% for calls and 101% for puts in mean Gain Ratio [11].

Another important advantage is data efficiency. Learning the residual with three years of data can achieve hedging performance comparable to directly learning the full hedging function with ten years of data [11]. This matters because clean financial data can be difficult to obtain, and market regimes can change over time. A structure-guided model that needs less data is often more practical than a purely black-box model.

This evidence changes how Black-Scholes should be interpreted. The model is not simply wrong because its assumptions are unrealistic. Its delta still contains valuable economic information, and deep learning can be used to correct the remaining error. However, this evidence focuses on hedging rather than direct pricing accuracy, so Gain Ratio should not be directly compared with pricing MSE. The approach also requires high-quality option Greeks and implied-volatility data, which may be harder to obtain in less liquid markets.

5. Comparative discussion and future development

The development across the three research stages is clear. The first stage establishes a theoretical benchmark. The second stage improves pricing accuracy by using deep learning. The third stage improves hedging performance by combining Black-Scholes delta with residual learning. These

studies should not be forced into one single evaluation metric. A theory paper should be judged by its contribution to the no-arbitrage pricing logic. A pricing model should be judged mainly by pricing errors. A hedging model should be judged by hedging errors and risk reduction.

A key limitation in this comparison is that the two empirical studies use different versions of the Black-Scholes benchmark. The pricing study evaluates Black-Scholes with historical volatility estimated from annualized 20-day intraday returns, which partly explains its relatively high MSE. The hedging study uses implied Black-Scholes delta, where implied volatility already contains market expectations. Therefore, the pricing MSE improvement in Arin's research and the hedging Gain Ratio improvement in Qiao's research should not be read as a direct like-for-like comparison between a weak and a strong Black-Scholes model [8, 11]. They answer different questions under different volatility inputs (Table 1).

Table 1. Summary comparison of Black-Scholes and deep learning extensions

Reference group	Model role	Quantitative evidence	Future implication
[1-3]	Theoretical benchmark	No empirical MSE reported; provides closed-form price and delta	Use as an interpretable baseline, not as a data-driven predictor
[8]	Hybrid deep learning pricing model	Average pricing MSE reduced from 10.9518 to 0.59810; 94.5% improvement against a historical-volatility Black-Scholes benchmark	Test robustness across assets, liquidity levels, and market regimes
[11]	Residual deep learning hedging model	Gain Ratio improvement often exceeds 100%; 129% for calls and 101% for puts in overall gain ratio against direct learning, using implied Black-Scholes delta as benchmark	Develop structure-guided models with lower data requirements

The main future direction is structure-guided deep learning. Pure black-box models may fit option prices well in a particular dataset, but they are harder to interpret and may perform poorly when market conditions change. A stronger direction is to preserve economically meaningful components, such as the Black-Scholes price, delta, vega, and implied-volatility structure, while using neural networks to learn residual errors. This design keeps part of the model interpretable while still allowing enough flexibility to capture nonlinear market behaviour.

A second direction is to include realistic market frictions. Deep hedging research has already begun to consider transaction costs, market impact, liquidity constraints, and risk limits [6]. These factors are important because real hedging is never perfectly continuous or costless. Newer studies also investigate neural stochastic differential equation models and implied-volatility-surface feedback for hedging option books [12, 13]. These approaches are promising because they view the option market as a dynamic system rather than as a set of isolated contracts.

A third direction is uncertainty-aware modelling. A useful model should not only produce a price or hedge ratio; it should also communicate how reliable that output is. Recent research on deep gamma hedging and uncertainty-aware deep hedging suggests that model disagreement, moneyness, volatility regimes, and market stress can all affect hedging reliability [14, 15]. In practical trading, a slightly less aggressive hedge with clearer uncertainty information may be better than a highly optimized hedge that fails without warning. Future research should therefore report not only average errors, but also tail losses, CVaR, stress-period performance, and results by maturity and moneyness group.

Overall, the literature suggests that future option-pricing research should move beyond the simple debate between formulas and neural networks. The strongest models will likely combine no-

arbitrage restrictions, volatility-surface information, market microstructure variables, and transparent performance metrics.

6. Conclusion

This review examined how the Black-Scholes model has developed from a classical theoretical benchmark into a foundation for deep learning extensions in option pricing and hedging. The Black-Scholes model remains essential because it introduced the no-arbitrage logic that still supports modern derivative valuation. Its strength lies in its structure, interpretability, and ability to generate a clear benchmark price and delta. However, its assumptions of constant volatility, frictionless markets, continuous hedging, and lognormal asset prices are too restrictive for many real market conditions. These assumptions can lead to systematic pricing and hedging errors, especially when volatility smiles, liquidity differences, discrete trading, and market frictions become important.

Recent deep learning research addresses these weaknesses in two main ways. Pricing-focused models use neural networks to reduce the difference between predicted option prices and observed market prices. The evidence reviewed in this paper shows a substantial quantitative improvement: a hybrid classifier-pricer model reduces average MSE from 10.9518 under a historical-volatility Black-Scholes benchmark to 0.59810, which represents a 94.5% improvement. Hedging-focused models use a different evaluation logic. Instead of minimizing pricing error, they aim to reduce mean squared local hedging error. The residual-learning approach is especially convincing because it does not discard Black-Scholes delta. It learns a correction term on top of the implied Black-Scholes delta and often improves hedging performance by more than 100% in Gain Ratio terms. However, these two sets of results should be interpreted separately because the first uses historical volatility in Black-Scholes pricing, while the second uses implied Black-Scholes delta as the hedging benchmark. The overall implication is still consistent: deep learning is most useful when it corrects a clearly defined Black-Scholes baseline rather than replacing financial structure completely.

The main conclusion is that Black-Scholes should not be treated as an outdated model that must be completely replaced. A more productive view is that Black-Scholes provides the economic skeleton, while deep learning adds flexible correction ability. The future of option pricing is likely to be hybrid: interpretable no-arbitrage structures combined with data-driven learning, realistic market frictions, volatility-surface information, and uncertainty measurement. For future studies, stronger out-of-sample testing, cross-market datasets, explainability methods, and stress-period evaluation will be necessary. These improvements can help ensure that deep learning models are not only accurate in backtests, but also reliable enough for real trading, hedging, and risk management.

References

- [1] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 637–654.
- [2] Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1), 141–183.
- [3] Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3), 229–263.
- [4] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2), 327–343.
- [5] Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *The Journal of Finance*, 49(3), 851–889.
- [6] Buehler, H., Gonon, L., Teichmann, J., & Wood, B. (2019). Deep hedging. *Quantitative Finance*, 19(8), 1271–1291.

- [7] Ruf, J., & Wang, W. (2020). Neural networks for option pricing and hedging: A literature review. *Journal of Computational Finance*, 24(1), 1–46.
- [8] Arin, E., & Ozbayoglu, A. M. (2022). Deep learning based hybrid computational intelligence models for options pricing. *Computational Economics*, 59, 39–58.
- [9] Ruf, J., & Wang, W. (2022). Hedging with linear regressions and neural networks. *Journal of Business & Economic Statistics*, 40(4), 1442–1454.
- [10] Chen, J., & Li, L. (2021). Data-driven hedging of stock index options via deep learning. arXiv:2111.03477.
- [11] Qiao, C., & Wan, X. (2024). Enhancing Black-Scholes delta hedging via deep learning. arXiv:2407.19367.
- [12] Cohen, S. N., Reisinger, C., & Wang, S. (2022). Hedging option books using neural-SDE market models. arXiv:2205.15991.
- [13] Francois, P., Gauthier, G., Godin, F., & Perez Mendoza, C. O. (2024). Enhancing deep hedging of options with implied volatility surface feedback information. arXiv:2407.21138.
- [14] Armstrong, J., & Tatlow, G. (2024). Deep gamma hedging. arXiv:2409.13567.
- [15] Poddar, M. (2026). Uncertainty-aware deep hedging. arXiv:2603.10137.