

Hull-White Model Calibration: Neural Network Beats Traditional Method on Speed, Accuracy, and Stability

Benedikt Grimus

November 2025

Introduction

Financial institutions rely on mathematical models to price interest rate derivatives and manage risk. A critical step in this process is calibration, where a model's parameters are adjusted to accurately reflect current market conditions. Traditional calibration methods are computationally intensive, often requiring several minutes to run, and can produce, depending on the initial conditions, unstable results. In today's fast-paced financial environment, this computational bottleneck limits the ability to perform real-time risk analysis and pre-trade pricing. This research addresses this challenge by investigating whether a Neural Network (NN), can provide a faster, more accurate, and more reliable alternative to these established techniques for calibrating the widely used one-factor Hull-White (HW) interest rate model.

Methodology

The study conducted a direct comparison between a custom-developed NN and the industry-standard Levenberg-Marquardt (LM) optimization algorithm. To ensure a robust and fair evaluation, three different implementations of the LM method were tested, each using a distinct strategy for initialization. The core task for each method was to calibrate the HW model using a comprehensive dataset of European at-the-money swaptions from the Euro market, covering the period from June to August 2025.

A key innovation of this research is the NN's end-to-end training approach. Instead of being trained to simply replicate the outputs of the traditional optimizer, the NN was trained to directly minimize the actual pricing error between model-implied values and market-observed volatilities. The performance of all methods was assessed using a daily intra-day hold-out protocol, which evaluates each model's ability to price European swaptions that were not part of the calibration set, providing a true measure of out-of-sample predictive power.

Key Findings & Results

The empirical results demonstrate a clear and multifaceted superiority of the NN approach over traditional methods across all key performance criteria.

Superior Accuracy: The NN consistently produced more accurate pricing. It achieved an average out-of-sample error of 4.33 basis points (bps), an improvement over the 6.05 bps error

from the best-performing traditional LM method. The NN's predictions were also more consistent, with a day-to-day error variation approximately 4.6 times lower than the most stable LM strategy.

Transformational Speed: The NN offers a fundamental shift in computational efficiency. A full daily calibration was executed in approximately 4 milliseconds on average. In contrast, the most efficient LM method required over 72 seconds. This represents a speed-up factor of approximately 17,000, effectively transforming the calibration process from a time-consuming batch task into a near-instantaneous one.

Enhanced Stability and Reliability: The analysis of the model parameters over time revealed that the NN generated stable and economically coherent outputs that responded logically to market fluctuations and simulated stress scenarios. The traditional LM methods, in contrast, demonstrated varying degrees of instability. While naive initialization strategies produced either highly erratic parameters or exhibited a systematic parameter drift leading to performance degradation, the more sophisticated strategy proved considerably more robust. Nevertheless, even this best-performing traditional method yielded parameters that were substantially more volatile than those generated by the NN, confirming the superior reliability of the machine learning approach.

Implications & Recommendations

The findings of this thesis have practical implications for financial modeling, risk management, and trading operations.

Enablement of Real-Time Applications: The sub-second performance of the NN makes high-frequency applications operationally feasible. This includes real-time portfolio risk valuation, pre-trade pricing for large books of derivatives, and intraday recalibration in response to changing market conditions, capabilities that are unattainable with traditional methods.

Improved Risk Management: The higher accuracy and stability of the NN-based calibration provide a more reliable foundation for pricing and hedging derivatives. The model's ability to avoid the systematic biases and parameter drift seen in traditional methods leads to more trustworthy risk metrics and reduces model risk.

Strategic Recommendation: Financial institutions are encouraged to explore and adopt end-to-end trained machine learning frameworks for model calibration. The research demonstrates that this approach, which trains a model to directly optimize for a core business objective (i.e., minimizing pricing error), can unlock a step-change in performance. The significant upfront investment in data preparation and model training is justified by substantial long-term gains in speed, accuracy, and operational efficiency, providing a distinct competitive advantage.