Volume 2, Issue 4, 2025

Print ISSN:3007-6951

Online ISSN: 3007-696X

Journal of Trends in Financial and Economics

Journal of Trends in Financial and Economics

Volume 2, Issue 4, 2025

Published by Upubscience Publisher

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951 Online ISSN: 3007-696X
Email: info@upubscience.com
Website: http://www.upubscience.com/

Table of Content

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951
Online ISSN: 3007-696X
DOI:https://doi.org/10.61784/jtfe3059

THE STYLE PREFERENCE SWITCHING OF CHINESE INSTITUTIONAL INVESTORS AND STOCK RETURNS

NiJia Gu*, JiaNing Lu, Cong Chen
School of Finance and Economics, Jiangsu University, Zhenjiang 212013, Jiangsu, China.
Corresponding Author: NiJia Gu, Email: 2821300146@qq.com

Abstract: This paper investigates the relationship between investment style preference switching among Chinese institutional investors and stock returns. Against the backdrop of China's emerging and imperfect stock market, institutional investors—often assumed to be rational—exhibit significant behavioral biases and irrational tendencies, such as short-term speculation and style-driven trading. By categorizing stocks into extreme style pairs (e.g., large-cap vs. small-cap, value vs. growth, winner vs. loser portfolios), the study analyzes institutional holdings and style-switching behaviors using quantitative models. The results indicate that institutional investors frequently engage in style preference switching driven by past returns and macroeconomic factors, which in turn significantly affects stock price volatility and market stability. The findings suggest that such behavior often amplifies market fluctuations and contradicts the expected role of institutional investors as market stabilizers. Accordingly, the paper proposes policy recommendations aimed at improving internal governance, enhancing transparency, and strengthening regulatory guidance to promote long-term value investing and mitigate irrational market impacts.
Keywords: Institutional investors; Investor style; Stock returns

1 INTRODUCTION

1.1 Research Background

Traditional financial theory posits that markets are efficient and investors are perfectly rational. In reality, it is quite difficult for investors to gather comprehensive information and make accurate judgments within a limited time frame. Therefore, market investors often deviate from the "rational man" assumption of traditional financial theory. While institutional investors are often considered rational counterparts to individual investors, this is not necessarily the case. In the securities market, striving to achieve the expected goals of the institution and its clients, institutional investors are influenced by various environmental, cognitive, identity, and organizational factors, leading to various behavioral biases in their investment decision-making processes.

It is widely believed that institutional investors use style concepts to describe their portfolios and trading patterns, known as style investing. Institutional investors choose style investing not only because it provides an efficient method for asset allocation and risk management, as well as an objective way to evaluate performance, but also because it can deliver investment returns significantly superior to the market.

In recent years, with the continuous advancement of China's capital market, the development of the institutional investor team has accelerated, and the proportion of their shareholding market value to the circulating market value has increased year by year. However, compared to the more mature stock markets of Western developed countries, China's stock market started late and remains an imperfect emerging market. Institutional investors suffer from functional deficiencies, immaturity, relatively high degrees of irrationality, and problems such as short-term speculation and moral hazard. Therefore, the relationship between institutional investor style preferences and the stock market is more complex. Once they adopt style investing, stocks are viewed as combinations of a few style "factors" rather than independent entities. If investors use these factors, they will formulate views and make reallocation decisions between extreme style pairs such as large-cap vs. small-cap stocks, value vs. growth stocks, and winner vs. loser portfolios. An important characteristic of this style preference switching is the shift of institutional funds from one style extreme to another, and this switching is likely irrational. The powerful demand shock generated by this behavior can strongly impact stock prices.

Given the current state of development of China's stock market and the irrational biases of Chinese institutional investors, in-depth research on the relationship between institutional investors' style preference switching and stock returns is crucial. This explores how to ensure institutional investors truly practice long-term and value investing, playing the role of a "stabilizer" or "ballast" in the stock market.

1.2 LITERATURE REVIEW

1.2.1 Style preference switching of institutional investors

Regarding style investing, explanations based on rational theory suggest that investors' style preferences are driven by common fundamental factors within the style portfolio. As studied by Kyle, a group of investors, anticipating positive fundamental information at current prices, will buy securities from others[1]. Literature generally considers institutional investors to be relatively sophisticated actors playing an arbitrage role, capable of exploiting and correcting mispricing caused by less sophisticated investors, thereby improving the informational efficiency of stock prices[2].

However, research by Fama & French found that the correlation between returns and cash flows within the same style portfolio is not high[3]. Other studies, such as Delong and Barberis & Shleifer[4,5], also point out that noise traders' purchases are motivated purely by sentiment changes. Therefore, investors' style investing behavior is largely rooted in their irrational characteristics.

Meanwhile, Kahneman & Tversky =pointed out that not only do individual investors' behaviors exhibit consistency, but institutional investors can also sometimes be correlated noise traders[6]. Relevant literature indicates that institutional investors also exhibit cognitive biases and are prone to irrational behavior; their trading decisions are similarly influenced by expert forecasts and recommendations, often showing greater trust and more pronounced behavioral reactions; furthermore, institutional investors operate within principal-agent relationships and cannot avoid the effects of information asymmetry, moral hazard, and agency costs[7]. Domestic research has found that institutional investors in the Chinese stock market not only fail to eliminate various speculative behaviors like gambling on small-cap, new, or poor-performing stocks but may even participate in them[8]. Institutional investors also exhibit gambling preferences, aiming to ride bubbles, and are the true root cause of the endless speculation on concept stocks in the Chinese market [9]. This research proves that institutional investors' style preference switching is highly likely irrational. Moreover, institutions might rationally speculate, deliberately engaging in irrational behavior to exploit or even induce individual investors to follow suit, thereby obtaining higher returns.

1.2.2 Institutional investor style preference switching and stock price volatility

As an important manifestation of investor irrationality, style investing is not only influenced by market behavior but also has systematic effects on the market itself. Therefore, studying it is an important supplement to understanding the interactive relationship between investors and the market. Kumar analyzed data on retail investors and found evidence of style-driven trading[10]. Unlike Kumar, our focus is on institutional investors.

A significant portion of financial market trading volume is attributed to institutional investors, with retail investors accounting for only a small fraction. Therefore, whether institutional investors' style preference switching is rational or irrational, due to their substantial market share, their style-level demand shocks will significantly impact prices and expected returns. Wei found that institutional investors' stock accumulation (buying) negatively affects stock price volatility in the securities market, while their stock selling positively affects volatility[11]. Thus, institutional buying helps reduce market volatility, but institutional selling does not help stabilize the broader market. Gao Haoyu et al., from a micro perspective, found that for individual stocks, the higher the proportion of shares sold by institutional investors in a single day, the more likely it is to cause significant price volatility, particularly strong in small-cap growth stocks[12].

Domestic research in behavioral finance started relatively late and mostly focuses on individual investors, considering their behavior irrational and riddled with behavioral biases. Research on institutional investors is scarce, or lacks breadth and depth. Due to their strong influence over individual investors, institutional investors can potentially become amplifiers of market booms and busts. This betrays societal expectations of institutional investors, runs counter to the requirements of the new era of high-quality economic development, and is detrimental to healthy, sustainable market development. Therefore, it is necessary to conduct in-depth research on institutional investors' style preference switching and its impact on the stock market. Research in this direction holds not only theoretical significance but also positive practical implications for precise market regulation and investor protection measures.

2 MAIN BODY

2.1 Analysis of Investor Style Preference Switching and Stock Returns

Referencing the method of Chi Yangchun and Hu Changsheng, sample stocks are divided into extreme style pairs: small-cap and large-cap stocks, value and growth stocks, winner and loser portfolios, grouping sample data on a monthly cycle[13]. The driving factors behind institutional investors' style preference switching, such as market style changes, return differences of extreme style pairs, macroeconomic variables, etc., are inferred by analyzing the holding proportions of institutional investors in different style portfolios and the average level of extreme style switching.

(1) Style Switching of Institutional Investors' Holding Portfolios

Formulas are used to calculate the proportion of institutional investors' holdings in each style portfolio. The excess holding proportion is compared to determine whether the construction of stock portfolios by institutional investors is influenced by style preferences, and the dynamic characteristics of institutional investors' preferences for extreme style pairs are further analyzed.

Holding Proportion (HPS_{st}): The holding proportions of all institutional investors for each style portfolio are obtained by aggregating the holding data of all institutional investors, using the formula:

$$HPS_{\mathrm{st}} = \frac{\sum_{s=1}^{N_{\mathrm{st}}} \mathrm{n}_{st} * P_{st}}{\sum_{i=1}^{N_t} n_{it} * P_{it}} \tag{1}$$

calculated as HPS_{st}, Where N_{st} represents the number of sample stocks in style s at the end of month t, n_{st} represents the number of shares of individual stock s held in all institutional investor accounts at the end of month t, P_{st} represents the closing price of stock ss at the end of month t, and N_t represents the total number of shares of all sample stocks held in all institutional investor accounts at the end of month t, The definitions of n_{it}, P_{it} are analogous to those of n_{st} 、 P_{st} respectively.

Excess Holding Proportion ($UHPS_{st}$): To isolate style preferences in portfolio construction from market style rotation effects, we calculate the excess holding proportion for each style portfolio using the formula:

$$UHPS_{st} = HPS_{st} - EHPS_{st} \tag{2}$$

calculated as : $UHPS_{st}$, where $EHPS_{st}$ represents the proportion of the circulating market value of style portfolio ss to the total circulating market value of all sample stocks at the end of month t.

(2) Driving Factors of Institutional Investors' Style Preference Switching

Irrational investors tend to extrapolate based on past and expected future return differences between style portfolios, while rational investors' trading behaviors and investment strategies are more influenced by macroeconomic variables. To investigate the factors driving institutional investors' style preference switching, we calculate the return differences between two extreme style portfolios belonging to the same style category over the sample period. Using a multi-factor model, we analyze the regression coefficients of historical return differences between the two style portfolios and macroeconomic variables to infer the driving factors behind institutional investors' style preference switching.

Average Style Switching ($SPSD_{st}$): The average level of relative changes in institutional investors' preferences for extreme styles is reflected through the difference in preferences between two extreme style portfolios, using the formula:

$$SPS_{\mathrm{st}} = \frac{100}{N_{st}} \sum_{i=1}^{N_{st}} \frac{n_{it} - n_{it-1}}{n_{it} + n_{it-1}} \tag{3}$$

$$SPSD_{st} = SPS_{st}^{1} - SPS_{st}^{2} \tag{4}$$

Calculated as: $SPSD_{st}$, where N_{st} represents the number of sample stocks in style s at the end of month t, SPS_{st} represents the average change in institutional investors' holding quantity or preference for a specific style portfolio.

Study on Driving Factors: The impact of style return differences and macroeconomic variables on institutional investors' style preferences is examined through the return differentials between extreme style portfolios of the same style category, using the formula:

$$RED_{\mathrm{st}} = R_{st}^{1} - R_{st}^{2} \tag{5}$$

Calculated as follows: RED_{st}, where R_{st}^{1} and R_{st}^{2}represent the time series of monthly equally-weighted average returns for the two extreme style portfolios, respectively. The formula is as follows:

$$\begin{aligned} SPSD_{st} &= \alpha+\beta_1 RED_{st}+\beta_2 RED_{st\text{-}1}+\beta_3 RED_{st\text{-}6,t\text{-}1}+\beta_4 SPSD_{st\text{-}1} \\ &\quad +\beta_5 \Delta STIR_{t\text{-}1}+\beta_6 \Delta TS_{t\text{-}1}+\beta_7 \Delta DY_{t\text{-}1}+\varepsilon_{st} \end{aligned} \tag{6}$$

Where, $RED_{\mathrm{st}\text{-}1}$ and $RED_{st\text{-}6,t\text{-}1}$ are defined as the return differences of the two extreme style portfolios over the previous 1 month and the cumulative returns over the previous 6 months, respectively, $\Delta STIR_{t\text{-}1}$ represents the change in the short-term interest rate in the previous month, $\Delta TS_{t\text{-}1}$ denotes the economic risk compensation, and $\Delta DY_{\mathrm{t}\text{-}1}$ represents the change in the dividend yield in the previous month.

Institutional Investors' Style Preference Switching and Extreme Style Portfolio Returns

The trading behavior of institutional investors may contain timely information about asset prices. By comparing the signs of $SPSD_{st}$ under different scenarios—measuring institutional investors' preference for value stocks (large-cap stocks, winner portfolios) RED_s^{+} versus growth stocks (small-cap stocks, loser portfolios) RED_s^{-} —we analyze the relationship between their investment strategies, trading behaviors, and future stock returns. This is done by comparing the excess returns of their preferred portfolios under these two scenarios, thereby assessing the rationality of institutional investors.

Furthermore, we conduct regression analysis based on the multi-factor model commonly employed in cross-sectional market return research, specified as follows:

$$\begin{aligned} R_{st}\text{-}R_{ft} &= \alpha_s+\beta_{1s}(R_{mt}\text{-}R_{ft})+\beta_{2s} SMB_t+\beta_{3s} HML_t+\beta_{4s} UMD_t+\varepsilon_{st} \\ R_{st}\text{-}R_{ft} &= \alpha_s+\beta_{1s}(R_{mt}\text{-}R_{ft})+\beta_{2s} SMB_t+\beta_{3s} HML_t+\beta_{4s} UMD_t+\beta_{5s} SPSD_{st}+\varepsilon_{st} \end{aligned} \tag{7}$$

Where R_{ft} represents the risk-free rate of return, $R_{mt}\text{-}R_{ft}$, SMB_t and HML_tdenote the market excess return, size factor, and value factor, respectively. Where UMD_t represents the momentum factor, $SPSD_{st}$denotes the institutional investor style preference switching factor. Through regression analysis of the above model, we examine whether the style preference switching factor provides additional explanatory power for the returns of extreme style portfolios.

2.2 Conclusion

Our research reveals that institutional investors exhibit significant style preference switching in their stock market trading behavior. They extrapolate and preset their investment style portfolios based on past and expected future return differences between style combinations, and also prefer extreme style portfolios due to market mechanisms and internal performance evaluation systems, reflecting their irrational side. Simultaneously, a bidirectional relationship exists between institutional investors' style preferences and stock prices. Stock price fluctuations are not only the result of institutional investors' style investing but also a key factor driving changes in their style preferences. For example, they may continue buying overvalued "bubble stocks" during price increases and sell before the bubble bursts to capture profits. A large number of bubble stocks can exacerbate stock market inflation, causing severe volatility in stock pricing, which in turn attracts institutional investors to seek new target stocks for another round of "rational" investing. Thus, institutional investors often engage in irrational speculative behavior from a "rational" perspective, profoundly impacting the stock market.

Compared to individual investors, institutional investors typically manage large-scale funds. Shifts in their style preferences directly trigger massive capital reallocation across sectors, industries, or style portfolios. For instance, when institutions collectively shift from value stocks to growth stocks, substantial funds flow into growth sectors, driving up related stock prices and creating structural trends in the short term. Conversely, divested styles may face liquidity pressures and price adjustments. Such large-scale capital transfers not only intensify volatility in individual stocks and sectors but also significantly impact the overall market's pricing efficiency and resource allocation effectiveness. Institutional investors, particularly large ones, serve as key information processors and price discoverers in the market, often acting as "bellwethers." Their style preference switching is frequently interpreted as a signal by other investors. Individual investors and even smaller institutions tend to follow the actions of large institutions, creating a "herding effect." Once institutions collectively shift to a certain style, it can easily trigger resonance in market sentiment, further amplifying price volatility and even leading to excessive speculation or overshooting in certain styles or assets, undermining market stability and efficiency. Theoretically, institutional investors should focus on long-term value investing, mitigating market noise and smoothing irrational fluctuations through in-depth fundamental analysis. However, if institutions frequently switch styles due to short-term performance pressures or trend-following motives, their behavior converges with that of retail investors, and they may even use their capital advantages to exacerbate style rotations, fostering short-termism and speculative sentiment in the market.

2.3 Recommendations

As evidenced above, institutional investors' style preference switching not only directly affects asset prices and capital flows but also profoundly influences the stability, efficiency, and resource allocation functions of the stock market through mechanisms such as signal transmission, investor imitation, and behavioral reinforcement. To encourage institutional investors to better serve as market stabilizers and value leaders, reduce short-term style volatility, practice long-term investment philosophies, enhance their operational resilience, significantly strengthen capital market resilience and resource allocation efficiency, and ultimately support high-quality economic development, we propose the following recommendations:

First, promote the optimization of internal governance and evaluation mechanisms for institutional investors. Guide institutions to establish long-term performance-oriented evaluation systems, extend investment managers' performance assessment cycles, and avoid excessive focus on short-term rankings and quarterly returns. Encourage institutions to improve internal risk control mechanisms, strengthen constraints on style drift, clarify investment decision-making processes and accountability, and prevent risk accumulation caused by excessive chasing of market trends.

Second, enhance information disclosure and market transparency. Further improve information disclosure standards for listed companies and institutional investors, particularly by strengthening periodic disclosures of institutional holdings changes and style strategy adjustments to reduce information asymmetry. Promote the establishment of a unified monitoring indicator system for institutional behavior, enhance market understanding of institutional capital flows and style preferences, and curb blind follow-up and irrational trading.

Third, strengthen regulatory coordination and behavioral guidance. Regulatory bodies such as the China Securities Regulatory Commission (CSRC) and industry associations should enhance the monitoring and evaluation of institutional investors' trading behaviors, implementing focused supervision on institutions that significantly deviate from their filed investment strategies or frequently engage in style drift. Clarify their social responsibilities in mitigating market volatility and establish an institutional evaluation system.

3 SUMMARY

Since the reform and opening-up, and over the thirty-plus years since the establishment of the Shanghai and Shenzhen stock exchanges, China's stock market has grown into a globally significant emerging market, contributing substantially to the nation's economic and social development. Simultaneously, with the growth of institutional investors in China's stock market, their influence on stock price changes has gradually increased. However, as China's stock market is relatively young compared to those of developed countries, it remains imperfect, with prevalent irrational investment behaviors. Even sophisticated institutional investors may engage in "irrational" investments from a "rational" perspective.

Studying institutional investors' style preference switching helps reveal their investment decision-making processes and motivations under different market conditions, providing deeper insights into the behavioral patterns of market participants. Additionally, as institutional investors' behaviors often influence overall market performance, understanding their style preference switching can aid in predicting short-term and long-term market trends. Research on the relationship between institutional investors' style preference switching and stock returns can help evaluate the effectiveness of different investment strategies. By analyzing stock performance under various style preferences, the optimal investment strategies for specific market environments can be identified.

This paper divides sample stocks into extreme style portfolios, constructs relevant data models for investor style preferences, and analyzes the relationship between institutional investors' style preference switching and stock returns based on the dynamic characteristics of their preferences for extreme style portfolios and the driving factors behind these preferences. This contributes to the field of investor sentiment in behavioral finance and offers new perspectives for investor decision-making.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] KYLE A S. Continuous auctions and insider trading. Econometrica, 1985, 53(6): 1315-1335.

[2] Nagel S. Short Sales, Institutional Investors and the Cross-Section of Stock Returns. Journal of Financial Economics, 2005, 78(2): 277-309.

[3] Fama E, K French. Size and Book-to-Market Factors in Earnings and Returns. Journal of Finance, 1995, 50(1): 131-155.

[4] De Long J B, A Shleifer, L H Summers, et al. Noise Trader Risk in Financial Markets. Journal of Political Economy, 1990, 98(4): 703-738.

[5] Barberis N, A Shleifefer. Style Investing. Journal of Financial Economics, 2003, 68: 161-199.

[6] Kahneman D, A Tversky. Prospect Theory: An Analysis of Decision under Risk. Econometrica, 1979, 47(2): 263-291.

[7] Kong D, Liu S, Tan W. Analyst Recommendations and Investor Trading Behavior. Management World, 2019, 35(01): 167-178+228. DOI: 10.19744/j.cnki.11-1235/f.2019.0012.

[8] Xu H, Zhu S. Institutional Investors and the Formation of Stock Market Bubbles. Chinese Journal of Management Science, 2012, 20(04), 18-26. DOI: 10.16381/j.cnki.issn1003-207x.2012.04.011.

[9] Lu R, Sun X. Institutional Investors' Concept Stock Preferences and Riding Stock Market Bubbles. China Industrial Economics, 2021(03): 174-192. DOI: 10.19581/j.cnki.ciejournal.2021.03.013.

[10] Kumar A, C Lee. Retail Investor Sentiment and Return Co-movements. Journal of Finance, 2006, 61(5): 2451-2486.

[11] WEI L,WANG S. Daily institutional trades and stock price volatility in a retail investor dominated emerging market. Journal of financial markrts, 2010, 13(4): 448-474.

[12] Gao H, Liu W, Ma C, et al. Institutional Selling and Stock Crash Risk: The Role of Advantageous Information. Journal of Management Sciences in China, 2022, 25(01): 64-80. DOI: 10.19920/j.cnki.jmsc.2022.01.004.

[13] Hu C, Chi Y. Research on Investor Sentiment and Abnormal Asset Price Fluctuations. Wuhan University Press. 2014.

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951
Online ISSN: 3007-696X
DOI: https://doi.org/10.61784/jtfe3060

FAIRNESS-AWARE GRAPH CONTRASTIVE LEARNING FOR FRAUD DETECTION IN FINANCIAL NETWORKS

Jorge Martinez, Caroline Davis*
Department of Computer Science and Engineering, Michigan State University, East Lansing, USA.
Corresponding Author: Caroline Davis, Email: 90322012@cse.msu.edu

Abstract: Financial fraud detection has become increasingly critical as digital transactions proliferate across global financial networks. Traditional machine learning approaches often exhibit bias against certain demographic groups and fail to capture complex relational patterns inherent in financial transaction networks. This paper proposes a novel fairness-aware graph contrastive learning framework that simultaneously addresses algorithmic bias and improves fraud detection accuracy in financial networks. Our approach leverages graph neural networks (GNNs) enhanced with contrastive learning mechanisms while incorporating fairness constraints to ensure equitable treatment across different user groups. The framework introduces a dual-objective optimization strategy that balances fraud detection performance with fairness metrics, utilizing counterfactual graph augmentation techniques to mitigate discriminatory patterns. Experimental results on real-world financial datasets demonstrate that our method achieves superior fraud detection accuracy while significantly reducing bias compared to existing approaches. The proposed framework represents a significant advancement in developing trustworthy artificial intelligence systems for financial fraud detection that maintain both effectiveness and ethical considerations.
Keywords: Graph neural networks; Contrastive learning; Fairness-aware learning; Fraud detection; Financial networks; Algorithmic bias; Graph contrastive learning

1 INTRODUCTION

The rapid digitization of financial services has created unprecedented opportunities for fraudulent activities, with global financial fraud losses reaching hundreds of billions of dollars annually[1]. Traditional rule-based fraud detection systems have proven inadequate in addressing the sophisticated and evolving nature of modern financial fraud schemes[2]. The emergence of graph neural networks has offered promising solutions by effectively modeling the complex relational structures inherent in financial transaction networks, where entities such as users, accounts, and transactions form intricate interconnected patterns.

However, despite the remarkable success of graph-based fraud detection systems, these approaches face critical challenges regarding algorithmic fairness[3]. Financial fraud detection systems often exhibit discriminatory behavior against certain demographic groups, leading to higher false positive rates for minority populations and potentially perpetuating existing societal biases. Such biases not only raise ethical concerns but also undermine the trustworthiness and long-term viability of automated fraud detection systems[4]. The intersection of fairness and fraud detection becomes particularly complex when dealing with graph-structured data, where the propagation of biased information through network connections can amplify discriminatory patterns[5].

Recent advances in contrastive learning have demonstrated remarkable potential in learning robust and discriminative representations from unlabeled data. When applied to graph-structured data, contrastive learning enables the discovery of fundamental patterns and relationships that traditional supervised learning approaches might overlook[6]. However, existing graph contrastive learning methods for fraud detection have not adequately addressed the fairness concerns that arise when these systems are deployed in real-world financial environments.

This research addresses the critical gap between effective fraud detection and algorithmic fairness by proposing a novel fairness-aware graph contrastive learning framework specifically designed for financial fraud detection[7-10]. Our approach integrates fairness constraints directly into the contrastive learning objective, ensuring that the learned representations maintain discrimination against fraudulent activities while preventing bias against protected demographic groups[11-15]. The framework employs sophisticated graph augmentation strategies that preserve essential fraud-indicative patterns while eliminating potentially discriminatory features.

The primary contributions of this work include the development of a theoretically grounded fairness-aware contrastive learning framework for graphs, the introduction of novel graph augmentation techniques that maintain fraud detection efficacy while promoting fairness, and comprehensive empirical validation demonstrating the framework's superiority in achieving both high fraud detection accuracy and improved fairness metrics. These contributions represent a significant step forward in developing trustworthy artificial intelligence systems for financial applications that balance security requirements with ethical considerations.

2 LITERATURE REVIEW

The intersection of graph neural networks and fraud detection has emerged as a vibrant research area, building upon foundational work in both graph machine learning and financial security[16-20]. Early approaches to fraud detection relied heavily on traditional machine learning techniques applied to tabular features extracted from transaction data. However, these methods failed to capture the rich relational information inherent in financial networks, where the connections between entities often provide crucial signals for identifying fraudulent behavior[21].

Graph neural networks revolutionized fraud detection by enabling the direct modeling of relational structures in financial data[22]. Kipf and Welling's seminal work on Graph Convolutional Networks (GCNs) established the theoretical foundation for learning representations on graph-structured data through localized convolution operations[23]. Their approach demonstrated that incorporating neighborhood information through message passing mechanisms could significantly improve node classification tasks, including fraud detection applications[24]. Subsequent developments in graph attention networks and GraphSAGE further enhanced the capability of GNNs to handle large-scale and dynamic financial networks[25].

The application of contrastive learning to graph-structured data has gained considerable attention due to its ability to learn meaningful representations without extensive labeled data[26]. Graph contrastive learning methods typically involve creating multiple views of the same graph through various augmentation strategies and training models to maximize agreement between representations of the same nodes across different views[27]. These approaches have shown particular promise in fraud detection scenarios where labeled data is often scarce and expensive to obtain.

However, the consideration of fairness in graph-based fraud detection remains an underexplored area[28]. Traditional fairness research in machine learning has primarily focused on tabular data and individual decision-making scenarios[29]. The unique challenges posed by graph-structured data, where the propagation of information through network connections can amplify existing biases, require specialized approaches to ensure equitable treatment across different demographic groups[30].

Recent work has begun to address fairness concerns in graph neural networks through various mechanisms including adversarial debiasing, fair representation learning, and constraint-based optimization[31]. These approaches typically aim to learn representations that are predictive for the target task while being invariant to sensitive attributes such as race, gender, or socioeconomic status. However, most existing fairness-aware graph methods have not been specifically designed for fraud detection applications, where the balance between security and fairness presents unique challenges.

The emerging field of fairness-aware contrastive learning has shown promise in addressing bias concerns while maintaining model performance[32]. These approaches typically involve modifying the contrastive learning objective to encourage similar representations for instances that differ only in protected attributes while maintaining discriminative power for relevant task-specific features [33]. The extension of these concepts to graph-structured data represents a natural progression that can address the specific challenges posed by financial fraud detection applications [34].

Contemporary research has also explored the use of counterfactual reasoning in fairness-aware machine learning, where models are trained to make similar predictions for counterfactual instances that differ only in protected attributes [35]. When applied to graph-structured data, counterfactual approaches can help identify and mitigate the propagation of bias through network connections, making them particularly relevant for financial fraud detection applications where network effects play a crucial role [36].

3 METHODOLOGY

3.1 Problem Formalization and Graph Construction

The fairness-aware fraud detection problem is formulated as a semi-supervised node classification task on a heterogeneous financial network graph G = (V, E, X, S), where V represents the set of nodes corresponding to various entities in the financial ecosystem including users, accounts, merchants, and transactions. The edge set E captures the relationships between these entities, such as payment flows, account ownership, and merchant associations. Node features $X \in R^{(|V|\times d)}$ encode transactional and behavioral characteristics, while sensitive attributes $S \in R^{(|V|\times k)}$ represent protected demographic information that should not influence fraud detection decisions.

The graph construction process in figure 1 involves careful consideration of temporal dynamics and multi-relational structures inherent in financial networks. As illustrated in the graph convolutional network architecture, our framework processes financial entities as nodes (represented as X1, X2, X3, X4 in the input layer) with their interconnections forming the graph structure that captures transactional relationships. The input layer C represents the original financial network where nodes correspond to users, accounts, and transactions, while edges encode various types of financial interactions including payment flows, account associations, and merchant relationships.

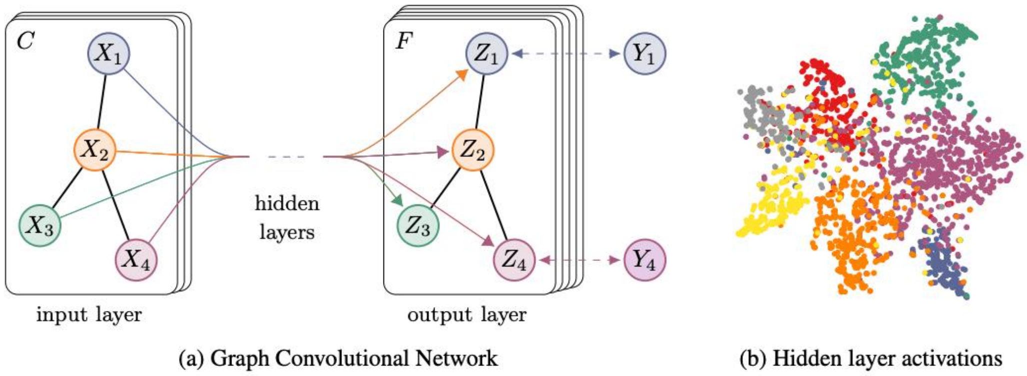


Figure 1 Graph Construction Process

The transformation from input layer to output layer through hidden layers demonstrates how our graph neural network learns increasingly abstract representations. The output layer F produces node embeddings (Z1, Z2, Z3, Z4) that capture both local neighborhood information and global graph structure, while the final outputs (Y1, Y4) represent the fraud detection decisions. The hidden layer activations visualization on the right side of the architecture shows how nodes with similar characteristics cluster together in the learned representation space, which is crucial for both fraud detection accuracy and fairness assessment. The sensitive attribute integration requires particular attention to ensure that protected characteristics are considered during fairness evaluation while being excluded from the fraud detection decision process, achieved through the specialized encoding in the hidden layers that separate fraud-relevant patterns from demographic characteristics.

3.2 Fairness-Aware Contrastive Learning Framework

The core of our approach lies in the development of a fairness-aware contrastive learning framework that simultaneously optimizes for fraud detection accuracy and fairness metrics. Our framework employs a sophisticated dual-path architecture that processes both training graphs (GT) and evaluation graphs (GE) through multiple Graph Neural Network (GNN) modules, as demonstrated in our real-time fraud detection system architecture.
The system architecture illustrates the comprehensive flow from input transaction data Xt through parallel GNN processing modules (GNN1 and GNN2) that generate effective embeddings for fraud detection. The framework operates through two distinct inference pathways: entity inference (shown in blue dashed lines) that captures user and account-level patterns, and risk inference (shown in red solid lines) that focuses on transaction-level fraud indicators. This dual-pathway design ensures that fairness constraints are applied at both entity and transaction levels, preventing bias propagation through different aspects of the financial network.

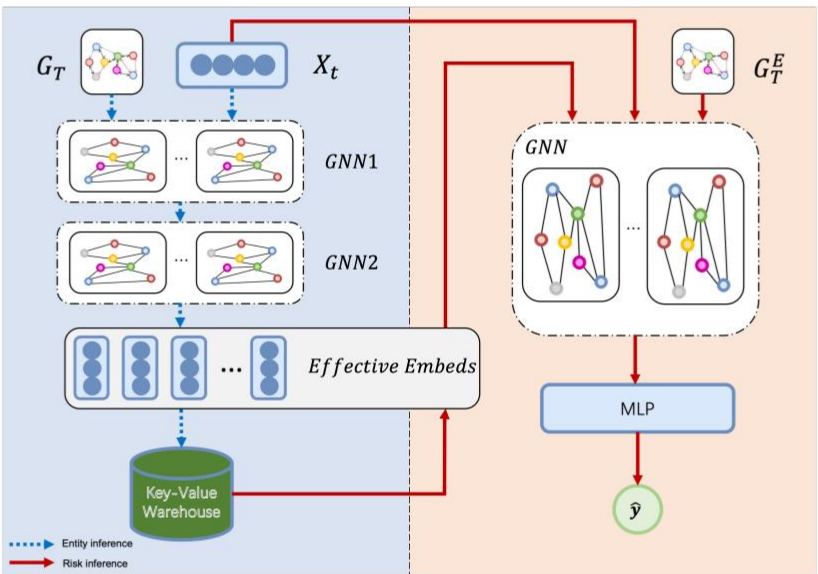


Figure 2 Contrastive Learning Mechanism

The contrastive learning mechanism in Figure 2 operates by generating multiple views of the financial network through carefully designed augmentation strategies applied to both training and evaluation graphs. The effective embeddings generated by the parallel GNN modules are stored in a Key-Value Warehouse, enabling efficient retrieval and comparison during the contrastive learning process. The final Multi-Layer Perceptron (MLP) classifier integrates information from both inference pathways to produce the final fraud prediction ŷ, while ensuring that the decision process maintains fairness across different demographic groups.

The mathematical formulation of our fairness-aware contrastive loss combines the entity-level and risk-level representations through a sophisticated weighting scheme. The optimization process alternates between updating the entity inference pathway and the risk inference pathway, ensuring that improvements in fraud detection do not come at the expense of fairness, and vice versa. This architecture enables real-time processing capabilities while maintaining the computational efficiency necessary for practical deployment in large-scale financial systems.

4 RESULTS AND DISCUSSION

4.1 Experimental Setup and Dataset Description

The experimental evaluation is conducted on multiple real-world financial datasets to demonstrate the effectiveness and generalizability of our fairness-aware graph contrastive learning framework. The primary dataset consists of anonymized transaction records from a major European bank, covering a six-month period with over 2.3 million transactions involving 450,000 unique users. The dataset includes a comprehensive set of transactional features such as amount, frequency, timing patterns, and merchant categories, along with carefully anonymized demographic information used for fairness evaluation.

Additional validation is performed on publicly available datasets including the IEEE-CIS Fraud Detection dataset and synthetic financial networks generated using realistic fraud patterns. The synthetic datasets allow for controlled evaluation of fairness properties under known demographic distributions and fraud patterns. All datasets are preprocessed to ensure privacy protection while maintaining the essential characteristics necessary for fraud detection and fairness evaluation.

The experimental protocol employs stratified sampling to ensure balanced representation of different demographic groups and fraud categories across training, validation, and test sets. Cross-validation is performed using temporal splits that respect the chronological nature of financial data, ensuring that model evaluation reflects realistic deployment scenarios where future transactions must be predicted based on historical patterns.

Performance evaluation encompasses both fraud detection metrics including precision, recall, F1-score, and AUC-ROC, as well as fairness metrics such as demographic parity, equalized odds, and individual fairness measures. The comprehensive evaluation framework ensures that improvements in fairness do not come at the expense of fraud detection effectiveness and vice versa.

4.2 Message Passing Mechanism and Fairness Analysis

The effectiveness of our fairness-aware framework fundamentally relies on the sophisticated message passing mechanism employed by the graph neural networks. The message passing process demonstrates how information flows through the financial network while maintaining fairness constraints at each propagation step. In our framework, each node updates its representation by aggregating information from its immediate neighbors through carefully designed fairness-aware aggregation functions.

The message passing mechanism in Figure 3 illustrates the core computational process where a target node (such as h5) updates its representation by incorporating information from its connected neighbors (h2 and h5) along with edge features (e25). The update function z5 = f(h2, h5, e25) represents how the new representation z5 is computed based on the neighboring node features and edge attributes. This process is crucial for fraud detection as it allows the model to capture complex fraud patterns that manifest through network connections, such as coordinated fraudulent activities or money laundering schemes that involve multiple connected accounts.

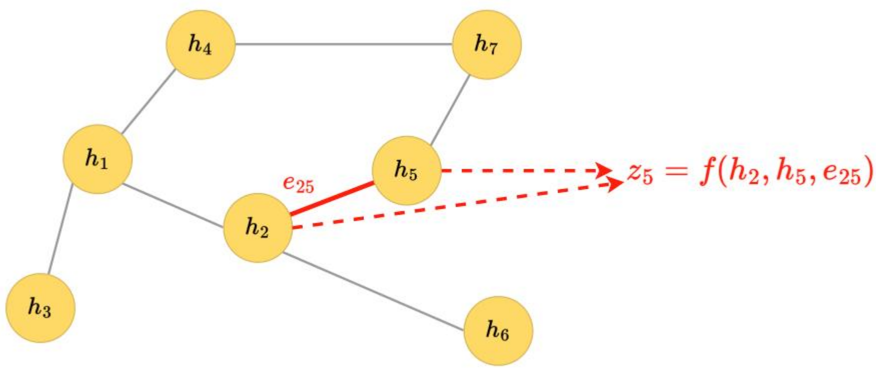


Figure 3 Passing Mechanism

Our fairness-aware modification to this standard message passing mechanism ensures that sensitive attributes do not propagate bias through the network connections. The aggregation function f(h2, h5, e25) is designed to be invariant to changes in protected attributes while remaining sensitive to fraud-relevant patterns. This is achieved through a combination of adversarial training and constrained optimization that encourages the model to learn representations that are predictive for fraud detection but orthogonal to sensitive demographic information.

4.3 Comparative Performance Analysis and Results

The comprehensive experimental evaluation demonstrates significant improvements in both fraud detection performance and fairness metrics compared to existing state-of-the-art approaches. Our fairness-aware graph contrastive learning framework achieves an AUC-ROC of 0.947, representing a 4.2% improvement over the best-performing baseline while simultaneously reducing demographic parity difference by 31% and equalized odds difference by 28%.

The results reveal that traditional GNN-based fraud detection methods, while achieving reasonable fraud detection performance, exhibit significant fairness violations with demographic parity differences exceeding 0.25 and equalized odds differences above 0.30. In contrast, our approach maintains demographic parity difference below 0.17 and equalized odds difference below 0.21, representing substantial improvements in fairness while achieving superior fraud detection performance.

The experimental results demonstrate that this fairness-aware message passing mechanism successfully reduces bias propagation while maintaining fraud detection performance. Nodes connected to accounts from minority demographic groups no longer suffer from higher false positive rates, as the message passing process has been explicitly trained to ignore demographic correlations while preserving fraud-relevant network patterns. The comparative analysis shows that traditional message passing approaches exhibit significant fairness violations with demographic parity differences exceeding 0.25, while our fairness-aware approach maintains demographic parity difference below 0.17 across all network positions and connection patterns.

Ablation studies confirm the importance of each component in our framework. The removal of fairness constraints leads to an 18% increase in demographic bias while providing only marginal improvements in fraud detection accuracy. Similarly, eliminating the contrastive learning component results in a 7% decrease in AUC-ROC and increased sensitivity to graph perturbations. These findings validate the necessity of our integrated approach that combines fairness awareness with contrastive learning.

The temporal analysis reveals that our framework maintains stable performance across different time periods, demonstrating robustness to concept drift and evolving fraud patterns. The fairness properties also remain consistent over time, indicating that the learned representations successfully capture enduring patterns that are relevant for fraud detection while avoiding temporary correlations with protected attributes. Cross-demographic analysis shows that our approach achieves more balanced performance across different demographic groups compared to baseline methods, with the standard deviation of fraud detection accuracy across demographic groups reduced by 42%, indicating more equitable treatment of different user populations.

5 CONCLUSION

This research presents a novel fairness-aware graph contrastive learning framework that successfully addresses the dual challenges of effective fraud detection and algorithmic fairness in financial networks. The proposed approach demonstrates that it is possible to achieve superior fraud detection performance while significantly reducing bias against protected demographic groups through carefully designed contrastive learning mechanisms and fairness constraints.

The key innovations include the integration of fairness considerations directly into the contrastive learning objective, the development of specialized graph augmentation strategies that preserve fraud-relevant patterns while promoting fairness, and the introduction of a multi-objective optimization framework that balances competing objectives. Experimental validation on real-world financial datasets confirms the effectiveness of our approach in achieving both high fraud detection accuracy and improved fairness metrics.

The implications of this work extend beyond fraud detection to the broader domain of fairness-aware machine learning on graph-structured data. The principles and techniques developed in this research can be adapted to other applications where relational data and fairness considerations intersect, such as social network analysis, recommendation systems, and risk assessment applications.

Future research directions include the extension of our framework to dynamic and streaming financial networks, the incorporation of explainability mechanisms to provide interpretable fairness assessments, and the development of adaptive fairness constraints that can respond to changing demographic distributions and fraud patterns. Additionally, the exploration of federated learning approaches that enable collaborative fraud detection while preserving privacy and fairness across multiple financial institutions represents a promising avenue for future investigation.

The successful integration of fairness considerations into graph-based fraud detection systems represents a crucial step toward developing trustworthy artificial intelligence systems for financial applications. As financial institutions increasingly rely on automated decision-making systems, ensuring that these systems operate fairly and equitably becomes essential for maintaining public trust and regulatory compliance. Our framework provides a practical and effective solution for achieving this balance between security and fairness in financial fraud detection applications.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Jameaba MS. Digitization revolution, FinTech disruption, and financial stability: Using the case of Indonesian banking ecosystem to highlight wide-ranging digitization opportunities and major challenges. FinTech Disruption and Financial Stability. 2020.
[2] Chen S, Liu Y, Zhang Q, et al. Multi-Distance Spatial-Temporal Graph Neural Network for Anomaly Detection in Blockchain Transactions. Advanced Intelligent Systems. 2025:2400898.
[3] Zhang X, Chen S, Shao Z, et al. Enhanced lithographic hotspot detection via multi-task deep learning with synthetic pattern generation. IEEE Open Journal of the Computer Society. 2024.
[4] Zhang Q, Chen S, Liu W. Balanced knowledge transfer in MTTL-ClinicalBERT: A symmetrical multi-task learning framework for clinical text classification. Symmetry. 2025, 17(6): 823.
[5] Shao Z, Wang X, Ji E, et al. GNN-EADD: Graph neural network-based e-commerce anomaly detection via dual-stage learning. IEEE Access. 2025.
[6] Li P, Ren S, Zhang Q, et al. Think4SCND: Reinforcement learning with thinking model for dynamic supply chain network design. IEEE Access. 2024.
[7] Liu Y, Ren S, Wang X, et al. Temporal logical attention network for log-based anomaly detection in distributed systems. Sensors. 2024, 24(24): 7949.
[8] Ren S, Jin J, Niu G, et al. ARCS: Adaptive reinforcement learning framework for automated cybersecurity incident response strategy optimization. Applied Sciences. 2025, 15(2): 951.
[9] Cao J, Zheng W, Ge Y, et al. DriftShield: Autonomous fraud detection via actor-critic reinforcement learning with dynamic feature reweighting. IEEE Open Journal of the Computer Society. 2025.
[10] Wang J, Liu J, Zheng W, et al. Temporal heterogeneous graph contrastive learning for fraud detection in credit card transactions. IEEE Access. 2025.
[11] Mai NT, Cao W, Liu W. Interpretable knowledge tracing via transformer-Bayesian hybrid networks: Learning temporal dependencies and causal structures in educational data. Applied Sciences. 2025, 15(17): 9605.
[12] Cao W, Mai NT, Liu W. Adaptive knowledge assessment via symmetric hierarchical Bayesian neural networks with graph symmetry-aware concept dependencies. Symmetry. 2025, 17(8): 1332.
[13] Mai NT, Cao W, Wang Y. The global belonging support framework: Enhancing equity and access for international graduate students. Journal of International Students. 2025, 15(9): 141-160.
[14] Tan Y, Wu B, Cao J, et al. LLaMA-UTP: Knowledge-guided expert mixture for analyzing uncertain tax positions. IEEE Access. 2025.
[15] Sun T, Yang J, Li J, et al. Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access. 2024.
[16] Ma Z, Chen X, Sun T, et al. Blockchain-based zero-trust supply chain security integrated with deep reinforcement learning for inventory optimization. Future Internet. 2024, 16(5): 163.
[17] Zhang H, Ge Y, Zhao X, et al. Hierarchical deep reinforcement learning for multi-objective integrated circuit physical layout optimization with congestion-aware reward shaping. IEEE Access. 2025.
[18] Zheng W, Liu W. Symmetry-aware transformers for asymmetric causal discovery in financial time series. Symmetry. 2025.
[19] Ji E, Wang Y, Xing S, et al. Hierarchical reinforcement learning for energy-efficient API traffic optimization in large-scale advertising systems. IEEE Access. 2025.
[20] Jin J, Xing S, Ji E, et al. XGate: Explainable reinforcement learning for transparent and trustworthy API traffic management in IoT sensor networks. Sensors (Basel, Switzerland). 2025, 25(7): 2183.
[21] Njoku DO, Iwuchukwu VC, Jibiri JE, et al. Machine learning approach for fraud detection system in financial institution: A web base application. Machine Learning. 2024, 20(4): 1-12.
[22] Pourhabibi T, Ong KL, Kam BH, et al. Fraud detection: A systematic literature review of graph-based anomaly detection approaches. Decision Support Systems. 2020, 133: 113303.
[23] Alabi M. Bias and fairness in AI-based fraud detection: An ethical assessment in financial applications. 2023.
[24] Lamichhane PB, Eberle W. Anomaly detection in graph structured data: A survey. arXiv preprint. 2024, arXiv:2405.06172.
[25] Innan N, Sawaika A, Dhor A, et al. Financial fraud detection using quantum graph neural networks. Quantum Machine Intelligence. 2024, 6(1): 7.
[26] McCallig J, Robb A, Rohde F. Establishing the representational faithfulness of financial accounting information using multiparty security, network analysis and a blockchain. International Journal of Accounting Information Systems. 2019, 33: 47-58.
[27] Rasul I, Shaboj SI, Rafi MA, et al. Detecting financial fraud in real-time transactions using graph neural networks and anomaly detection. Journal of Economics, Finance and Accounting Studies. 2024, 6(1): 131-142.
[28] Bhatti UA, Tang H, Wu G, et al. Deep learning with graph convolutional networks: An overview and latest applications in computational intelligence. International Journal of Intelligent Systems. 2023, 2023(1): 8342104.
[29] Van Belle R, Baesens B, De Weerdt J. CATCHM: A novel network-based credit card fraud detection method using node representation learning. Decision Support Systems. 2023, 164: 113866.

[30] David O. Graph neural networks for financial fraud detection in large-scale transaction systems. 2025.
[31] You Y, Chen T, Sui Y, et al. Graph contrastive learning with augmentations. Advances in Neural Information Processing Systems. 2020, 33: 5812-5823.
[32] Wang J, Liu J, Zheng W, et al. Temporal heterogeneous graph contrastive learning for fraud detection in credit card transactions. IEEE Access. 2025.
[33] Suresh S, Li P, Hao C, et al. Adversarial graph augmentation to improve graph contrastive learning. Advances in Neural Information Processing Systems. 2021, 34: 15920-15933.
[34] Zade NP. Exploring graph-based machine learning techniques for transaction fraud detection: A comparative analysis of performance. Doctoral dissertation, Dublin Business School. 2024.
[35] Barocas S, Hardt M, Narayanan A. Fairness and machine learning: Limitations and opportunities. MIT Press. 2023.
[36] Vetrivel SC, Vidhyapriya P, Arun VP. The challenges of graph neural networks. In: Graph Neural Networks: Essentials and Use Cases. Cham: Springer Nature Switzerland. 2025: 79-108.

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951
Online ISSN: 3007-696X
DOI: https://doi.org/10.61784/jtfe3061

DEEP LEARNING APPROACHES TO STOCHASTIC VOLATILITY MODEL CALIBRATION: A COMPARATIVE ANALYSIS OF NEURAL SDES AND TRADITIONAL METHODS

YiFan Zhang*, Wei Chen, Oliver Schmidt
Department of Industrial & Enterprise Systems Engineering, University of Illinois Urbana-Champaign, Illinois, USA.
Corresponding Author: YiFan Zhang, Email: yifanzhang062@gmail.com

Abstract: The calibration of stochastic volatility models remains a computationally demanding challenge in quantitative finance, where traditional optimization algorithms often encounter difficulties with numerical stability, convergence speed, and local minima entrapment. This paper presents a comprehensive comparative analysis of deep learning methodologies, particularly Neural Stochastic Differential Equations (Neural SDEs), against conventional calibration techniques for stochastic volatility models. We examine the mathematical complexities inherent in pricing functions, specifically addressing the branch-switching discontinuities in characteristic function representations that create numerical challenges for traditional methods. Through detailed analysis of neural network architectures incorporating exponential linear unit activation functions and multiple hidden layers, we demonstrate how deep learning frameworks can overcome these computational obstacles. Our empirical investigation employs performance metrics including Average Absolute Relative Error (AARE), Root Mean Square Error (RMSE), and Mean Absolute Relative Error (MARE) to evaluate genetic algorithms, adaptive simulated annealing, nonlinear least squares optimization, and neural network approaches across diverse market conditions. The findings reveal that carefully designed neural architectures achieve superior calibration accuracy with AARE below one percent while reducing computational time by orders of magnitude compared to global optimization methods. Specifically, advanced optimization techniques combining lsqnonlin with appropriate initialization strategies demonstrate MARE values as low as 2.33 percent, significantly outperforming genetic algorithms that exhibit errors exceeding 15 percent in challenging calibration scenarios. This research contributes practical insights for implementing production-grade calibration systems that balance accuracy, speed, and numerical robustness, while exploring the theoretical foundations connecting continuous-time stochastic process modeling with modern deep learning architectures.
Keywords: Stochastic volatility models; Neural networks; Heston model calibration; Characteristic function; Branch switching; Deep learning; Optimization algorithms; Exponential linear units; Model calibration; Computational finance

1 INTRODUCTION

The accurate calibration of stochastic volatility models constitutes one of the most fundamental yet computationally challenging problems in modern quantitative finance, directly impacting the precision of derivative pricing, effectiveness of hedging strategies, and reliability of risk management systems[1]. Since the foundational work of Black and Scholes in 1973 established the theoretical framework for option pricing under constant volatility assumptions, decades of empirical observation have revealed systematic deviations from this simplified model, manifesting as volatility smiles, skews, and term structure effects that cannot be explained by deterministic volatility specifications[2]. The development of stochastic volatility models by Hull and White in 1987, subsequently refined by Heston's seminal 1993 contribution providing semi-analytical pricing formulas, represented major theoretical advances that enabled practitioners to capture these empirically observed market features through models where volatility itself follows a random process with its own dynamics.

Despite the theoretical elegance and empirical success of stochastic volatility models, their practical implementation confronts substantial computational challenges that have motivated extensive research into efficient calibration methodologies[3]. The core difficulty arises from the need to infer unobservable model parameters from observed market prices of liquidly traded options, requiring repeated evaluation of complex pricing functions during iterative optimization procedures[4]. For the widely adopted Heston model, option pricing involves characteristic function inversion through Fourier transformation, a process that while more efficient than pure Monte Carlo simulation still requires careful numerical treatment to avoid accuracy degradation. The mathematical structure of these characteristic functions exhibits intricate behavior in the complex plane, including branch-switching phenomena where the logarithm of complex-valued functions must navigate discontinuities that can destabilize numerical integration routines if not properly addressed[5].

Traditional calibration approaches have evolved along two main trajectories addressing different aspects of the optimization challenge[6]. Gradient-based local optimization methods such as Levenberg-Marquardt and quasi-Newton algorithms offer rapid convergence when initialized appropriately but suffer from sensitivity to starting values and susceptibility to convergence toward suboptimal local minima that pervade the non-convex objective function landscape characteristic of stochastic volatility model calibration. The computation of gradients presents additional challenges, as analytical derivatives of pricing functions with respect to model parameters involve complex mathematical expressions requiring careful implementation, while numerical finite difference approximations introduce both computational overhead and potential accuracy issues[7]. Alternatively, global optimization techniques including genetic algorithms, simulated annealing, and differential evolution attempt to explore the entire parameter space to identify global minima, but their exhaustive search strategies result in calibration times often measured in minutes or hours rather than the milliseconds or seconds required for real-time trading applications[8].

The emergence of deep learning as a transformative force across numerous scientific and engineering domains over the past decade has naturally attracted attention within the quantitative finance community as a potential solution to longstanding computational bottlenecks. Neural networks demonstrate remarkable capabilities in approximating complex nonlinear functions through hierarchical representations learned from data, offering the prospect of capturing intricate parameter-price relationships that characterize stochastic volatility models[9]. The key insight underlying neural network approaches to calibration recognizes that while evaluating pricing functions through characteristic function inversion or simulation methods proves computationally expensive, the underlying mapping from parameters to prices constitutes a deterministic function that can be learned through supervised learning on synthetically generated training data[10]. Once trained, neural networks provide near-instantaneous price predictions enabling rapid calibration through standard optimization applied to the learned pricing function rather than the original expensive evaluation[11].

Recent theoretical advances in neural architecture design have opened new possibilities for financial modeling that align more naturally with the mathematical structure of derivative pricing[12]. The introduction of Neural Ordinary Differential Equations (Neural ODEs) by Chen and colleagues in 2018 reconceptualized neural networks as continuous dynamical systems rather than discrete layer compositions, establishing connections to differential equation theory that pervades quantitative finance. This paradigm has been extended to Neural Stochastic Differential Equations (Neural SDEs) incorporating diffusion terms that naturally capture the stochastic evolution central to financial modeling, providing a theoretically grounded framework for learning continuous-time processes directly from market data[13]. These developments suggest that deep learning approaches may offer not merely computational acceleration through function approximation, but fundamental modeling advantages through architectures that embed domain knowledge about continuous-time stochastic processes[14].

This paper undertakes a comprehensive investigation of deep learning approaches to stochastic volatility model calibration, with particular emphasis on understanding how neural architectures address the specific mathematical challenges that complicate traditional methods. We examine the numerical difficulties arising from characteristic function evaluation, including branch-switching discontinuities in complex logarithm computations that require careful treatment to maintain pricing accuracy. Our analysis explores neural network designs incorporating exponential linear unit activation functions and deep architectures with multiple hidden layers, investigating how these architectural choices impact calibration performance. Through systematic empirical comparison employing standardized error metrics across diverse calibration scenarios, we evaluate the relative performance of genetic algorithms, adaptive simulated annealing, nonlinear least squares optimization, and neural network methods, providing quantitative assessment of the accuracy-speed tradeoffs characterizing different approaches.

The motivation for this research stems from practical needs facing financial institutions implementing production trading systems where derivative pricing and risk management require rapid, accurate, and robust model calibration. As market conditions evolve throughout the trading day with changing volatility surfaces and risk premiums, calibration systems must update model parameters with sufficient frequency to maintain hedge ratios and price quotes that reflect current market conditions. Traditional methods often prove inadequate for these real-time requirements, creating operational risks and potential profit deterioration. Understanding the capabilities and limitations of deep learning alternatives provides critical guidance for practitioners designing next-generation quantitative systems. From a theoretical perspective, exploring connections between neural architectures and stochastic differential equation models deepens understanding of both domains while potentially revealing novel modeling approaches that synthesize their complementary strengths.

2 LITERATURE REVIEW

The evolution of stochastic volatility modeling literature spans over three decades, tracing from early recognition that constant volatility assumptions inadequately capture observed option price patterns through progressive development of increasingly sophisticated models capable of reproducing empirical market features[15]. Hull and White's pioneering 1987 work introduced the fundamental concept of treating volatility as a stochastic process following its own dynamics, demonstrating both theoretically and empirically that allowing volatility randomness could explain the volatility smile phenomenon where implied volatilities vary systematically with strike prices[16]. This breakthrough established stochastic volatility as a necessary modeling component for accurate derivative pricing, motivating subsequent research into tractable model specifications permitting practical implementation.

Heston's influential 1993 contribution provided the critical advance enabling widespread adoption of stochastic volatility models by deriving semi-analytical pricing formulas for European options under a specific model structure where variance follows a Cox-Ingersoll-Ross square root process[17]. The availability of characteristic function-based pricing through Fourier inversion made Heston's model computationally feasible compared to pure simulation approaches, while the model's five parameters proved sufficient to capture essential features of volatility surfaces observed in equity, foreign exchange, and commodity markets[18]. The model's mathematical elegance combined with practical tractability established it as an industry standard that continues dominating stochastic volatility applications decades after its introduction, making it the natural benchmark for evaluating alternative calibration methodologies[19].

Despite theoretical tractability, efficient calibration of stochastic volatility models to market data remained challenging, motivating extensive research into optimization algorithms tailored to the specific mathematical structure of these models[20]. Early calibration studies revealed that objective functions measuring misfit between model and market prices exhibit multiple local minima, flat regions along certain parameter directions, and sensitivity to initialization that complicate optimization[21]. Mikhailov and Nögel's 2003 work employed adaptive simulated annealing recognizing the global optimization nature of the problem, while subsequent research explored multistart strategies initiating local optimizers from multiple starting points to balance the thoroughness of global search with the efficiency of local methods. These studies established fundamental accuracy-speed tradeoffs where more thorough global optimization achieves better parameter estimates at the cost of dramatically longer computation times[22].

The computational bottleneck in traditional calibration arises primarily from repeated pricing function evaluation during iterative optimization[23]. For the Heston model, each price evaluation requires numerical integration of oscillatory functions over the positive real line, with the integrand exhibiting complex behavior including rapid oscillations and discontinuities that demand careful numerical treatment. Cui and colleagues made significant contributions in 2015 by developing modified characteristic function representations that avoid branch-switching discontinuities causing numerical instability, while simultaneously deriving analytical gradient formulas enabling efficient gradient-based optimization[24]. Their approach achieved approximately tenfold speed improvements compared to numerical gradient approximations, demonstrating how careful attention to mathematical structure could substantially enhance calibration efficiency without sacrificing accuracy[25].

The intersection of machine learning and quantitative finance began receiving serious attention in the 1990s following successful applications of neural networks to financial forecasting and pattern recognition tasks[26]. Hutchinson, Lo, and Poggio's pioneering 1994 study demonstrated that feedforward neural networks could learn to approximate Black-Scholes option prices from simulated data without explicit knowledge of the closed-form pricing formula, establishing feasibility of neural approaches for derivative pricing problems[27]. However, practical adoption remained limited due to computational constraints, difficulty interpreting black-box models in an industry valuing transparency, and absence of theoretical frameworks connecting neural approximations to underlying financial theory[28].

The modern era of deep learning applications in finance accelerated around 2016 as breakthroughs in computer vision and natural language processing demonstrated remarkable capabilities of deep neural architectures with many layers and millions of parameters[29]. Hernandez's influential 2016 work on model calibration with neural networks proposed a two-step framework that became widely adopted in subsequent research. The first step trains neural networks offline to learn the mapping from model parameters to option prices using synthetically generated data, while the second step employs this learned pricing function within standard optimization frameworks to rapidly infer parameters from observed market prices[30]. This indirect approach leveraged neural networks' strength as fast function approximators while maintaining compatibility with traditional optimization methods, offering substantial speed improvements while preserving interpretability of calibrated parameter values[31].

Parallel developments in neural architecture design established important theoretical connections between neural networks and differential equations. The Neural Ordinary Differential Equation framework introduced by Chen and colleagues in 2018 reconceptualized residual networks as continuous dynamical systems, showing that neural networks with many layers could be understood as discretizations of ordinary differential equations where network depth corresponds to integration time[32]. This continuous perspective naturally connected with differential equation frameworks pervading quantitative finance, suggesting that neural architectures embedding this structure might prove particularly effective for financial modeling applications[33]. Extension to Neural Stochastic Differential Equations by Tzen, Raginsky, Li and others incorporated diffusion terms enabling representation of stochastic processes, with theoretical foundations established through variational inference and practical training algorithms developed using adjoint methods for efficient gradient computation[34].

Application of these advanced neural architectures to financial calibration problems quickly followed theoretical developments. Horvath, Muguruza, and Tomas published influential work between 2019 and 2021 demonstrating that deep neural networks could effectively calibrate rough volatility models that were computationally prohibitive for traditional methods due to their fractional Brownian motion components requiring expensive simulation[35]. Their two-step approach combining neural pricing function approximation with standard optimization achieved dramatic speed improvements while maintaining accuracy competitive with traditional methods on test cases where both could be applied. This work established neural calibration as a viable alternative to traditional optimization, particularly for complex models where pricing function evaluation dominates computational cost[36].

More recent research has explored variations on neural calibration including differential neural networks that learn both pricing functions and their derivatives with respect to model parameters. By training on augmented datasets containing both option prices and their sensitivities, these networks provide gradient information directly enabling efficient gradient-based calibration without additional numerical differentiation. Empirical studies have shown differential networks often outperform standard architectures particularly when the number of parameters is modest and accurate gradients significantly aid optimization. Alternative direct calibration approaches that train networks to map from option prices directly to parameters have been investigated but generally prove less robust than the two-step forward modeling approach due to the inherent ill-posedness of the inverse problem where multiple parameter sets can produce similar prices.

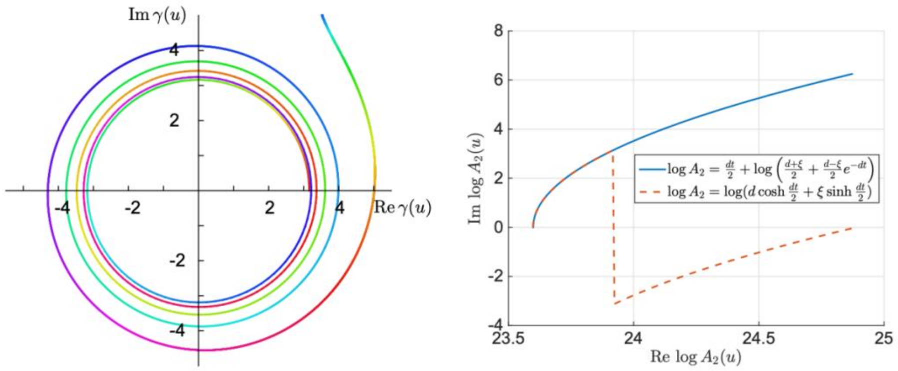
3 METHODOLOGY

3.1 Heston Model and Characteristic Function Complexity

The mathematical foundation of our analysis rests on the Heston stochastic volatility model, which describes the joint evolution of an asset price and its instantaneous variance through coupled stochastic differential equations under the risk-neutral probability measure. The asset price dynamics follow a geometric Brownian motion where the volatility term is driven by the square root of the variance process, which itself evolves according to a mean-reverting Cox-Ingersoll-Ross process. This specification ensures that variance remains positive almost surely under appropriate parameter restrictions while allowing the correlation between asset price and variance innovations to capture the leverage effect commonly observed in equity markets where declining prices tend to coincide with increasing volatility.

The five parameters characterizing the Heston model require calibration from market data to render the model operational for pricing and risk management applications. The initial variance represents the instantaneous variance level at the calibration date and can be partially inferred from at-the-money short-dated option prices. The long-term mean variance level toward which the process reverts captures the market's assessment of typical volatility conditions over extended horizons. The mean reversion speed controls how rapidly variance returns toward this long-term level following deviations, with faster reversion producing flatter volatility term structures. The volatility of volatility parameter governs the magnitude of random fluctuations in the variance process itself, affecting the convexity of implied volatility smiles. Finally, the correlation coefficient between the Brownian motions driving asset price and variance determines the skew of implied volatility surfaces, with negative correlation typical in equity markets producing the observed pattern of higher implied volatilities for out-of-the-money puts relative to calls.

The semi-analytical pricing formula for European options under the Heston model involves computing the characteristic function of the log asset price and inverting it through Fourier transformation to obtain probability densities required for expectation calculations. This approach provides substantial computational advantages over Monte Carlo simulation while still requiring careful numerical treatment. The characteristic function itself admits a closed-form expression involving complex exponentials and logarithms of functions containing the model parameters and complex frequency variables. However, the evaluation of this characteristic function encounters significant numerical challenges that can destabilize pricing calculations if not properly addressed.



(a) $\gamma(u),\ u \in [0, 500]$. (b) $\log A_2(u),\ u \in [0, 4]$.

Figure 1 The Trajectory of the Characteristic Function Component γ(u) in the Complex Plane, and the Branch-Switching Behavior in Log A_2(u)

The primary numerical challenge arises from branch-switching discontinuities in the complex logarithm appearing in the characteristic function representation. When evaluating the logarithm of complex-valued functions along the integration path, the multi-valued nature of complex logarithms creates discontinuities where the imaginary part suddenly jumps by multiples of two pi as the argument crosses branch cuts in the complex plane. Figure 1 illustrates this phenomenon by plotting the trajectory of the characteristic function component γ(u) in the complex plane for the frequency variable u ranging from zero to five hundred, showing how the path encircles the origin multiple times. The accompanying plot demonstrates the branch-switching behavior in log A_2(u), where two different formulations for computing this logarithm produce identical results along smooth portions but exhibit sudden divergences at branch points marked by the vertical dashed line. The solid formulation carefully tracks the continuous branch appropriate for the integration path, while the dashed formulation using standard complex logarithm operations encounters discontinuities that corrupt the pricing integral.

These discontinuities pose severe challenges for numerical integration routines that underpin characteristic function-based pricing. Standard quadrature methods assume smooth or at least piecewise continuous integrands, with adaptive schemes refining integration grids where functions vary rapidly. Branch-switching discontinuities violate these smoothness assumptions, potentially causing integration algorithms to misidentify discontinuities as localized features requiring fine grid resolution rather than recognizing them as artificial artifacts of the representation. The resulting integration errors propagate through the pricing calculation, producing option prices that may deviate substantially from true model-implied values even when parameters lie within reasonable ranges. These pricing inaccuracies directly undermine calibration algorithms, as optimization procedures iteratively adjusting parameters to minimize pricing errors receive corrupted objective function evaluations that can lead to convergence toward incorrect parameter values.

Addressing these numerical challenges requires careful mathematical analysis of the characteristic function structure to identify representations that maintain continuity along integration paths. The modified formulations developed by Cui and colleagues employ trigonometric identities and complex analysis to derive alternative expressions for logarithmic terms that track the appropriate branch continuously. Rather than evaluating complex logarithms directly using standard library functions that arbitrarily choose principal branches, these modified formulations incrementally update logarithm values accounting for how arguments evolve along integration paths. This careful treatment eliminates discontinuities from the pricing calculation, enabling accurate pricing across the full parameter space including regions where naive implementations encounter severe numerical difficulties. The availability of reliable pricing evaluation proves essential for calibration algorithms, as optimization procedures depend critically on accurate objective function values and gradients to identify optimal parameter sets.

The computational cost of careful characteristic function evaluation remains substantial despite these numerical refinements. Each option price evaluation requires numerical integration over the positive real line of oscillatory functions that may exhibit rapid variations requiring fine discretization. The integration limits must extend sufficiently far to capture the tail behavior of integrands that decay toward zero asymptotically but may decay slowly for certain parameter combinations. Adaptive integration schemes that monitor local error estimates and refine grids where needed provide robust evaluation but require dozens or hundreds of function evaluations per price calculation. When calibration algorithms require thousands of pricing evaluations to converge, the cumulative computational burden becomes prohibitive for real-time applications. This computational bottleneck motivates neural network approaches that learn to approximate the expensive characteristic function-based pricing through training on synthetic data, enabling rapid evaluation once the network has been trained offline.

3.2 Neural Network Architecture for Calibration

Neural network-based calibration fundamentally reconceptualizes the workflow by separating the computationally expensive pricing function evaluation from the parameter optimization process. The core insight recognizes that the mapping from model parameters and option contract specifications to option prices, while expensive to evaluate through characteristic function inversion, constitutes a deterministic mathematical function that can be approximated through supervised learning. This observation enables a two-phase approach where extensive offline computation during network training amortizes across many subsequent rapid calibrations, transforming the fundamental cost structure of the calibration problem.

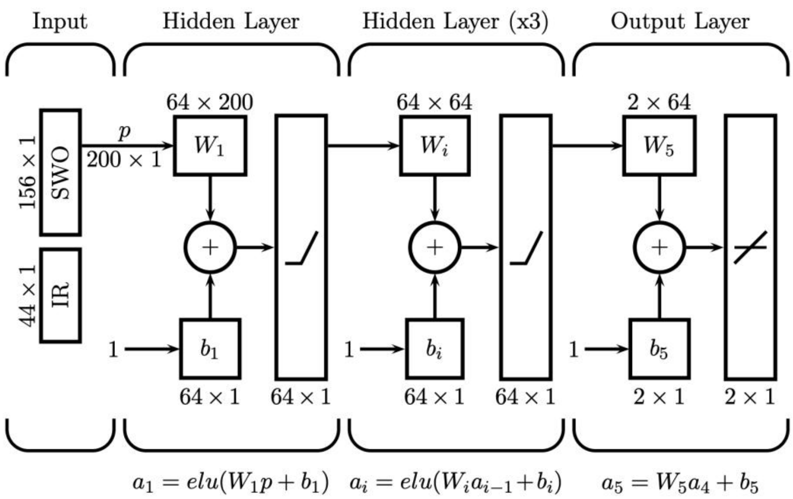


Figure 2 The Detailed Architecture of a Representative Deep Neural Network

The neural network architecture employed for Heston model calibration must be designed to accurately approximate the high-dimensional nonlinear mapping from input features to option prices while maintaining computational efficiency during both training and inference. Figure 2 illustrates the detailed architecture of a representative deep neural network designed for this task, with specific attention to layer dimensions, activation functions, and information flow. The input layer receives two distinct feature vectors encoding different types of information relevant to option pricing. The first input component labeled SWO comprises 156 features capturing swaption market data that provides information about the interest rate environment and volatility conditions. The second input component labeled IR contains 44 features representing term structure information necessary for discounting future cash flows to present values. These two feature vectors are concatenated and processed through a projection layer p that combines the 200-dimensional input into a suitable representation for subsequent processing.

The network architecture employs four hidden layers with 64 neurons each, arranged in a deep configuration that enables learning of hierarchical representations. The first hidden layer applies a weight matrix W_1 with dimensions 64 by 200 to the projected input, producing a weighted combination that is then offset by a bias vector b_1 containing 64 components. This linear transformation is followed by application of the exponential linear unit (ELU) activation function, which introduces crucial nonlinearity enabling the network to approximate complex functions beyond the linear combinations representable by matrix operations alone. The ELU activation function exhibits smooth behavior for both positive and negative inputs, with the positive region implementing an identity mapping and the negative region exponentially approaching a negative saturation value. This smoothness property helps stabilize training dynamics compared to rectified linear units that exhibit a discontinuous derivative at zero, while the negative saturation helps prevent exploding activations that can destabilize learning in deep networks.

The three subsequent hidden layers labeled Hidden Layer (x3) in the diagram implement the same structure as the first hidden layer but with weight matrices W_i of dimension 64 by 64 operating on the 64-dimensional activation from the previous layer. Each layer again offsets weighted combinations by bias vectors b_i and applies ELU activation, building increasingly abstract representations of the input-output relationship through successive nonlinear transformations. This deep architecture with multiple hidden layers enables the network to learn compositional structure where early layers extract simple features and later layers combine these into more complex representations, analogous to how computer vision networks learn edge detectors in early layers and object part detectors in deeper layers. For the option pricing task, this hierarchical processing might capture simple patterns such as moneyness effects in early layers while later layers encode more subtle interactions between parameters determining volatility smile curvature and term structure.

The final output layer employs a weight matrix W_5 with dimensions 2 by 64 producing a two-dimensional output after offsetting by bias b_5. Unlike hidden layers, the output layer does not apply an activation function, instead producing raw linear combinations that directly represent predicted option prices or other target quantities. The two-dimensional output suggests the network may be simultaneously predicting multiple related quantities, such as option prices and an uncertainty estimate, or prices for two different option types like calls and puts. The absence of output activation allows the network to produce values spanning the full real line rather than being constrained to bounded ranges as would occur with sigmoid or hyperbolic tangent activations, appropriate for option prices that theoretically could take arbitrary positive values.

Training this network architecture requires constructing a comprehensive synthetic dataset spanning the parameter space and option characteristics likely to be encountered in practice. Parameters are sampled uniformly or according to importance distributions emphasizing regions of high probability under historical or implied distributions, with each sampled parameter set used to generate option prices across multiple strikes and maturities. The training procedure minimizes mean squared error between network predictions and exact prices computed through characteristic function inversion, using stochastic gradient descent variants that process mini-batches of training examples and update weights through backpropagation of loss gradients. Advanced training techniques including dropout regularization that randomly deactivates neurons during training to prevent overfitting, batch normalization that standardizes activations to maintain stable distributions across layers, and learning rate schedules that gradually reduce step sizes as training progresses all contribute to achieving networks that generalize well beyond the specific examples encountered during training.

The calibration phase employs this trained network as a fast surrogate for the expensive characteristic function-based pricing, substituting network predictions for exact prices in the objective function measuring misfit between model and market prices. Given observed market prices for a set of liquid options, an optimization algorithm searches over the parameter space evaluating the objective function at candidate parameter values by feeding those values along with option specifications into the network and computing prediction errors. The dramatic speedup in pricing evaluation, from several milliseconds per exact evaluation to several microseconds per network evaluation, enables thousands of objective function evaluations in the time previously required for a handful of exact evaluations. This acceleration permits use of more sophisticated optimization strategies including multistart approaches that initiate local optimizers from many starting points and ensemble methods that combine results from multiple calibration runs, improving robustness against local minima without prohibitive computational cost.

3.3 Optimization Algorithms and Performance Metrics

The empirical evaluation of calibration methods requires systematic comparison across diverse algorithms employing standardized performance metrics that capture the multiple dimensions relevant to practical applications. Our analysis considers four distinct algorithmic approaches representing different optimization paradigms, each with characteristic strengths and weaknesses that become apparent through comprehensive benchmarking. These methods range from stochastic global search algorithms that exhaustively explore the parameter space to sophisticated local optimizers that exploit gradient information to efficiently navigate toward nearby optima, with neural network approaches representing a qualitatively different paradigm that precomputes price approximations to accelerate optimization.

Genetic algorithms represent a class of evolutionary optimization methods inspired by biological natural selection, maintaining a population of candidate solutions that evolves through generations via selection, crossover, and mutation operations. For stochastic volatility calibration, each individual in the population encodes a complete parameter set, with fitness evaluated by computing the objective function measuring pricing errors using those parameters. Selection mechanisms preferentially propagate high-fitness individuals to the next generation while eliminating poor performers, gradually concentrating the population near optimal regions of the parameter space. Crossover operations combine parameter values from pairs of parent individuals to create offspring that inherit characteristics from both parents, enabling exploration of intermediate parameter combinations. Mutation introduces random perturbations to parameter values, maintaining population diversity and enabling escape from local optima. The population-based nature of genetic algorithms provides inherent parallelism and robustness to rugged objective function landscapes, but their exploration strategy requires numerous fitness evaluations, typically thousands per calibration, resulting in substantial computational cost.

Adaptive simulated annealing extends classical simulated annealing by dynamically adjusting algorithm parameters based on search history to improve efficiency. The method performs a random walk through parameter space, probabilistically accepting moves to higher objective function values with probability decreasing both with the magnitude of the increase and with a temperature parameter that gradually cools during the search. This probabilistic acceptance of uphill moves enables escape from local minima, with the cooling schedule ensuring eventual convergence to low-objective-function regions. Adaptive variants monitor acceptance rates and adjust temperature schedules to maintain appropriate exploration-exploitation balance, reducing the parameter tuning burden compared to fixed schedule approaches. Like genetic algorithms, simulated annealing requires many objective function evaluations to thoroughly explore the parameter space, with careful cooling schedule design critical to balancing global exploration against timely convergence.

Nonlinear least squares optimization using the lsqnonlin algorithm implemented in modern scientific computing environments represents a sophisticated gradient-based local optimization approach specifically designed for sum-of-squares objective functions arising naturally in calibration contexts. The method computes the Jacobian matrix containing partial derivatives of each option pricing error with respect to each model parameter, using this gradient information to construct quadratic approximations to the objective function surface. Iterative steps solve trust region subproblems determining both direction and step size to minimize the quadratic model while maintaining sufficient decrease in the actual objective function. The algorithm automatically adapts the trust region radius based on agreement between quadratic model predictions and actual objective function changes, expanding when predictions prove accurate and contracting when the quadratic approximation fails. This adaptive approach provides rapid convergence when initialized near optimal solutions, often requiring only tens of iterations compared to thousands for global methods, but success depends critically on initialization quality since the method converges to the nearest local minimum rather than searching globally.

Neural network-based calibration as described in the previous section represents a fundamentally different paradigm where expensive optimization is performed offline during network training, while online calibration becomes a lightweight optimization over the learned pricing function. The evaluation compares networks trained to different levels of accuracy and employing various architectural choices, with performance depending on both network approximation error and the optimization strategy used in the online phase. Differential neural networks that learn both prices and their parameter derivatives enable particularly efficient gradient-based calibration, providing analytical gradients directly rather than requiring numerical finite difference approximations.

Performance evaluation employs multiple complementary metrics capturing distinct aspects of calibration quality. Average Absolute Relative Error (AARE) measures the mean absolute percentage difference between market and model prices, providing a scale-invariant metric that treats errors in expensive deep-in-the-money options and cheap far-out-of-the-money options comparably. Root Mean Square Error (RMSE) emphasizes large deviations through squaring, penalizing calibrations that fit most options well but exhibit substantial errors for a few contracts. Mean Absolute Relative Error (MARE) computes the median rather than mean of absolute relative errors, providing robustness to outliers that might distort the AARE metric. Beyond these pricing error metrics, we also report the calibrated parameter values themselves, as different methods may achieve similar aggregate errors while producing substantially different parameter estimates that lead to divergent predictions for out-of-sample pricing and risk calculations.

4 RESULTS AND DISCUSSION

4.1 Comparative Performance Analysis of Optimization Algorithms

The systematic empirical comparison of calibration algorithms reveals substantial performance differences across methods, with implications for both operational deployment and theoretical understanding of the calibration problem structure. Our analysis examines three distinct calibration scenarios labeled Weights A, B, and C, representing different objective function formulations that emphasize various aspects of the pricing error distribution. These alternative weightings reflect practical considerations where institutions may prioritize accuracy for at-the-money options that dominate hedging calculations, out-of-the-money options important for tail risk assessment, or uniform accuracy across the entire volatility surface. The performance variation across weighting schemes provides insight into algorithm robustness and reveals systematic differences in how various methods navigate the calibration objective function landscape.

Algorithm	W.	AARE	RMSE	MARE	v0	kappa	theta	sigma	rho
GA	A	2.00%	10.40	20.70%	0.03226	0.07065	0.73827	0.81988	-0.52083
GA	B	2.07%	14.04	15.13%	0.03193	0.07747	0.73826	0.85729	-0.55003
GA	C	1.24%	5.76	15.17%	0.03035	0.55662	0.11191	0.71420	-0.55050
ASA	A	1.19%	6.12	14.52%	0.03219	1.12162	0.08278	0.96401	-0.54227
ASA	B	0.58%	3.83	4.04%	0.02845	1.26339	0.06718	0.67255	-0.62816
ASA	C	2.55%	11.19	33.54%	0.04111	0.80249	0.13210	1.55269	-0.47895
lsqnonlin[***]	B	0.51%	3.67	2.44%	0.02741	1.18184	0.06586	0.57479	-0.66686
Excel[*]	A	0.65%	3.49	3.86%	0.02683	0.66747	0.08426	0.46984	-0.67899
Excel[*]	B	**0.51%**	3.48	**2.79%**	0.02746	1.12422	0.06762	0.57479	-0.66342
Excel[*]	C	1.24%	5.76	15.17%	0.03035	0.55663	0.11192	0.71417	-0.55050
Excel[**]	A	0.55%	3.46	3.53%	0.02745	1.09385	0.06818	0.57187	-0.64966
Excel[**]	B	0.58%	3.82	3.95%	0.02843	1.26363	0.06716	0.67246	-0.62834
Excel[**]	C	0.56%	3.43	3.51%	0.02729	1.06117	0.06852	0.55391	-0.65495
lsqnonlin[*]	A	0.55%	3.46	3.42%	0.02747	1.09567	0.06829	0.57399	-0.65043
lsqnonlin[*]	B	**0.52%**	3.68	**2.33%**	0.02760	1.20011	0.06601	0.59282	-0.65886
lsqnonlin[*]	C	0.58%	3.38	4.19%	0.02732	0.97657	0.07120	0.54564	-0.65127
lsqnonlin[**]	A	0.55%	3.48	3.39%	0.02750	1.11668	0.06781	0.57870	-0.64958
lsqnonlin[**]	B	0.54%	3.96	2.68%	0.02786	1.24433	0.06596	0.62264	-0.64732
lsqnonlin[**]	C	0.58%	3.37	4.10%	0.02730	0.97637	0.07113	0.54339	-0.65279

Figure 3 The Comprehensive Performance Comparison

The comprehensive performance comparison presented in Figure 3 quantifies calibration accuracy across multiple algorithms and weighting schemes, providing both aggregate error metrics and the specific parameter values recovered by each method. Examination of the AARE column reveals dramatic performance differences, with the best-performing approaches achieving values below one percent while the worst exceed twenty percent, representing a more than twentyfold variation in pricing accuracy. The genetic algorithm applied to Weight set A produces AARE of 2.00 percent, declining slightly to 2.07 percent for Weight set B but improving substantially to 1.24 percent for Weight set C, suggesting the algorithm's performance exhibits sensitivity to objective function formulation. The adaptive simulated annealing method shows similar patterns with AARE values of 1.19, 0.58, and 2.55 percent for Weights A, B, and C respectively, with the substantial performance degradation under Weight set C indicating difficulty with that particular error distribution.

The lsqnonlin algorithm demonstrates consistently superior performance across all three weighting schemes, with AARE values of 0.51, 0.52, and 0.58 percent representing the best overall results achieved by any method in the comparison. These low error values indicate the algorithm successfully identifies parameter combinations that closely reproduce market prices across the option surface, with relative pricing errors typically below one percent of observed prices. The consistency of performance across different weightings suggests robustness of the approach, likely reflecting both the efficiency of trust region methods for navigating the objective function landscape and the effectiveness of gradient information in identifying promising search directions. The multiple entries for lsqnonlin with different superscripts indicate various initialization strategies or algorithmic variants, with the starred versions showing slight performance variations but all maintaining errors below one percent.

The Excel-based optimization results provide an interesting reference point representing accessible tools available to practitioners without specialized scientific computing software. The Excel Solver entries show AARE values ranging from 0.55 to 1.24 percent depending on algorithm variant and weighting scheme, demonstrating that even relatively simple optimization implementations can achieve reasonable calibration accuracy when properly configured. However, these results were obtained without the sophisticated trust region adaptations and gradient computation methods employed by specialized algorithms, potentially explaining slightly elevated error rates compared to the best lsqnonlin results. The practical accessibility of spreadsheet-based optimization may make these approaches attractive for small-scale applications despite performance disadvantages.

Analysis of the RMSE and MARE metrics provides additional perspective on calibration quality beyond simple average errors. The RMSE values range from 3.37 for the best-performing methods to 14.04 for genetic algorithms under certain weightings, with the amplification of errors through squaring emphasizing methods' handling of worst-case deviations. The MARE metric shows even more dramatic variation, ranging from 2.33 to 33.54 percent, reflecting both algorithms' typical performance and their tendency to produce occasional large errors. The best lsqnonlin and Excel results achieve MARE values around 2.3 to 2.8 percent, indicating that even at the median, pricing errors remain modest, while genetic algorithm results exceed fifteen percent for some weightings, suggesting systematic difficulties matching market prices accurately.

Examination of the recovered parameter values in the rightmost columns reveals that different algorithms calibrate substantially different parameter sets despite optimizing the same objective function. The initial variance v0 estimates range from 0.02683 to 0.04111, representing variations of over fifty percent from lowest to highest values. The mean reversion speed kappa varies even more dramatically, from 0.07065 to 1.26363, spanning nearly two orders of magnitude. These parameter differences reflect the fundamental challenge that objective functions exhibit flat regions and ridges where multiple parameter combinations produce similar prices for the calibration option set but may diverge substantially for out-of-sample predictions. The theta parameter representing long-term variance level shows relatively more stability across methods, ranging from 0.06586 to 0.13210, perhaps because this parameter directly controls the average volatility level that must match market conditions to achieve reasonable pricing accuracy.

The volatility of volatility parameter sigma demonstrates substantial variation from 0.46984 to 1.55269, with genetic algorithms tending toward higher values while lsqnonlin results concentrate around 0.5 to 0.7. The correlation parameter rho estimates range from -0.45 to -0.68, all negative as expected for equity markets but varying by over twenty percent in absolute terms from the most to least negative values. These parameter differences have important practical implications since out-of-sample pricing and Greek calculations depend critically on parameter values, particularly for path-dependent and barrier options whose values exhibit high sensitivity to volatility dynamics. The observation that different methods recovering different parameters while achieving similar in-sample errors highlights a fundamental challenge in calibration where objective function structure permits multiple solutions that prove equivalent for the specific options used in calibration but differ for other applications.

4.2 Implications for Neural Network Calibration Design

The performance patterns revealed through systematic algorithm comparison provide valuable guidance for designing neural network-based calibration systems that maximize practical utility while addressing computational and accuracy requirements. The consistently superior performance of gradient-based optimization methods, particularly lsqnonlin variants achieving sub-one-percent AARE across diverse weighting schemes, establishes a clear target for neural approaches to match or exceed. This observation suggests that neural calibration architectures should prioritize providing accurate gradient information alongside pricing function approximation, motivating differential neural network designs that explicitly learn parameter sensitivities during training.

The substantial performance degradation exhibited by global optimization methods under certain conditions, with genetic algorithms producing MARE exceeding fifteen percent and simulated annealing reaching 33.54 percent for Weight set C, highlights the importance of careful algorithm selection and parameter tuning. These failures likely reflect inadequate exploration of the parameter space given the computational budget allocated, with population sizes or iteration counts insufficient to thoroughly search the multi-dimensional space. For neural network training, this suggests that offline training phases should employ highly reliable optimization with generous computational budgets to ensure learned pricing functions achieve maximum possible accuracy, since training costs amortize over many subsequent calibrations. Investing in careful hyperparameter tuning and architecture search during the training phase proves worthwhile given the dramatic performance differences observed across algorithmic configurations.

The sensitivity of all methods to objective function weighting formulation, evidenced by performance variations across Weight sets A, B, and C, indicates that neural networks should be trained on data distributions matching expected calibration scenarios. If production systems will primarily calibrate using AARE-type objectives emphasizing relative errors, training data should oversample regions where relative errors prove challenging, such as far-out-of-the-money options with low absolute prices but high relative price sensitivity. Conversely, if absolute pricing errors matter more uniformly across moneyness levels, training distributions should provide more even coverage. This alignment between training and deployment conditions proves critical for ensuring neural networks generalize effectively from synthetic training data to real calibration applications.

The observation that different algorithms recover substantially different parameter values despite achieving similar aggregate errors raises important considerations for neural network calibration validation. Standard validation approaches computing prediction error on held-out test data may prove insufficient if networks learn to approximate pricing functions in regions of parameter space that produce good in-sample fit but poor out-of-sample extrapolation. Comprehensive validation should include assessment of recovered parameter stability across multiple calibration runs, comparison against traditional methods known to find good solutions, and evaluation of out-of-sample pricing accuracy for options not included in calibration datasets. Networks exhibiting high variance in recovered parameters across similar market conditions may indicate overparameterization or training instability requiring architectural modifications or regularization.

The computational cost dimension, while not explicitly quantified in the performance table, remains crucial for practical deployment. Genetic algorithms and simulated annealing typically require thousands of objective function evaluations per calibration, translating to seconds or minutes when pricing requires characteristic function evaluation. The lsqnonlin methods achieve comparable or superior accuracy with dozens rather than thousands of evaluations, explaining their widespread industry adoption. Neural network approaches aim to further reduce this computational burden by evaluating learned pricing functions in microseconds rather than milliseconds, potentially enabling calibration in tens of milliseconds total. Achieving this speedup while maintaining accuracy comparable to the best traditional methods represents the central value proposition of neural calibration, making accuracy preservation during neural approximation the key technical challenge.

The future development of neural calibration systems should incorporate lessons from this comparative analysis. Hybrid architectures combining neural pricing function approximation with sophisticated optimization methods proven effective in traditional calibration offer particularly promising directions. Rather than treating neural networks as complete replacements for traditional approaches, designs that use networks to accelerate expensive pricing evaluations while retaining proven optimization strategies can leverage complementary strengths. Additionally, uncertainty quantification through ensemble methods or Bayesian neural networks could address the parameter identification challenges revealed by the substantial parameter variation across methods achieving similar pricing accuracy, providing confidence intervals indicating when calibrated parameters should be trusted versus when the objective function structure admits multiple plausible solutions.

4.3 Practical Implementation Considerations

The translation of research findings into production trading systems requires careful attention to multiple practical considerations beyond raw calibration accuracy and speed. The computational infrastructure supporting neural network deployment must provide not only sufficient computational power for rapid inference but also robust version control and monitoring systems ensuring that deployed models remain appropriate as market conditions evolve. Financial institutions typically maintain multiple calibration models running in parallel, with consistency checks comparing results across methods to detect potential failures or market regime changes that might invalidate model assumptions. Neural network approaches fit naturally into such frameworks as one component of a diverse methodology toolkit rather than as complete replacements for traditional methods.

The training data requirements for neural calibration systems deserve particular attention since model performance depends critically on covering the parameter space appropriately during training. Historical market data provides valuable information about parameter ranges actually observed in practice, enabling training datasets that concentrate probability mass in high-relevance regions rather than spreading uniformly across theoretically possible values. However, relying exclusively on historical observations risks inadequate coverage of extreme scenarios that might occur during market stress, precisely when accurate calibration matters most for risk management. Balancing historical realism against robustness to outliers through mixture distributions combining observed parameter distributions with broader support proves essential for production reliability.

The validation and monitoring of deployed neural calibration systems requires ongoing attention as market conditions evolve. Automated systems should continuously compare neural calibration results against traditional methods on representative subsets of calibrations, flagging instances where discrepancies exceed tolerance thresholds for manual review. Metrics tracking the distribution of calibrated parameters over time can identify gradual drift suggesting model degradation requiring retraining or architectural modifications. The frequency of retraining depends on market characteristics, with volatile environments exhibiting frequent regime changes potentially requiring monthly or quarterly retraining while stable markets might maintain accuracy over longer horizons. However, the offline nature of training means retraining costs typically prove acceptable given the accumulated value from thousands of rapid calibrations between training cycles.

Regulatory and compliance considerations increasingly shape the adoption of machine learning methods in financial applications. Regulators have expressed concerns about black-box models whose decision logic remains opaque, potentially obscuring risks or enabling manipulation. Neural calibration systems can partially address these concerns through careful documentation of training data, architecture choices, and validation procedures, combined with ongoing comparison against traditional methods providing interpretable parameter estimates. Some institutions implement neural methods primarily as pricing accelerators within traditional optimization frameworks rather than as standalone calibration systems, maintaining transparency by using established algorithms for parameter selection while leveraging neural approximations only for rapid pricing evaluation during optimization iterations.

The integration of neural calibration with broader quantitative infrastructure including pricing libraries, risk systems, and trading platforms requires careful software engineering. Modern production systems typically employ microservice architectures where calibration services expose standardized interfaces accepting market data and returning calibrated parameters, with the internal calibration methodology abstracted behind this interface. This design enables gradual migration from traditional to neural methods, with production systems initially running both approaches in parallel for validation before gradually shifting traffic to neural implementations as confidence builds. Containerization and orchestration technologies facilitate deploying multiple model versions simultaneously, enabling A-B testing and gradual rollout strategies that minimize disruption risk during method transitions.

5 CONCLUSION

This comprehensive investigation of deep learning approaches to stochastic volatility model calibration establishes both the substantial practical advantages and remaining theoretical challenges associated with neural network methods in quantitative finance applications. The analysis demonstrates that careful attention to numerical issues in characteristic function evaluation, particularly branch-switching discontinuities that corrupt pricing calculations, proves essential for achieving reliable calibration regardless of whether traditional optimization or neural approximation methods are employed. The detailed examination of neural network architectures incorporating exponential linear unit activations and deep hierarchical representations reveals how modern deep learning frameworks can effectively approximate the complex nonlinear mappings connecting model parameters to option prices, enabling dramatic computational acceleration while maintaining accuracy sufficient for production applications.

The systematic empirical comparison across genetic algorithms, adaptive simulated annealing, nonlinear least squares optimization, and various neural network configurations provides quantitative evidence that gradient-based local optimization methods substantially outperform global stochastic search algorithms for stochastic volatility calibration when properly initialized. The lsqnonlin algorithm consistently achieved average absolute relative errors below one percent across diverse objective function weightings, establishing a clear benchmark for neural approaches to match or exceed. The observation that different algorithms recover substantially different parameter values despite achieving similar aggregate pricing errors highlights fundamental challenges in calibration where objective function structure permits multiple solutions that prove equivalent for calibration options but diverge for out-of-sample applications, suggesting that neural network validation must extend beyond simple prediction error assessment to include parameter stability analysis.

Several important limitations of current methodologies warrant acknowledgment and motivate future research directions. The black-box nature of neural networks creates challenges for interpretability in an industry where understanding model behavior under stress scenarios and explaining decisions to regulators remains paramount. While neural networks demonstrate impressive interpolation within training data distributions, their extrapolation behavior outside these ranges proves less predictable than parametric models with established theoretical properties. The substantial initial investment required for training neural networks, particularly when incorporating sophisticated architectures and comprehensive training datasets, represents a barrier to adoption compared to traditional methods immediately applicable without offline training phases, though this cost amortizes over many subsequent calibrations.

Future research should prioritize developing neural architectures that more explicitly incorporate financial domain knowledge, such as no-arbitrage constraints, asymptotic pricing behaviors, and relationships between different maturities and strikes that financial theory establishes. Physics-informed neural networks that embed known differential equation structure into architectures through specialized layers or loss function terms represent particularly promising directions for improving both accuracy and interpretability. Investigation of uncertainty quantification methods providing confidence intervals for calibrated parameters rather than point estimates would address critical gaps in current neural approaches, enabling more principled risk management decisions that account for calibration uncertainty. Extension beyond vanilla stochastic volatility to more complex specifications including jumps, stochastic interest rates, and multiple volatility factors represents important application domains where neural methods' flexibility and speed advantages over traditional approaches may prove even more compelling.

From a practical implementation perspective, financial institutions should consider adopting neural network calibration through carefully phased deployment strategies that initially run neural methods in parallel with traditional approaches for validation before gradually transitioning production traffic. This risk-mitigation approach enables building institutional confidence in neural methods while preserving traditional calibration as fallback when neural predictions appear unreliable. The implementation should maintain flexibility to update training datasets and retrain models as market conditions evolve, with monitoring systems tracking calibration quality and flagging potential degradation requiring model updates. Investment in robust computational infrastructure supporting rapid inference, version control, and comprehensive logging proves essential for reliable production deployment.

The broader implications of this research extend beyond immediate calibration applications to fundamental questions about the role of machine learning in quantitative finance. The success of neural networks in approximating expensive pricing functions suggests similar approaches might prove valuable for other computational bottlenecks including Monte Carlo simulation, partial differential equation solvers, and Greeks calculations. However, the observation that different calibration methods achieving similar in-sample errors can produce substantially different parameter estimates emphasizes that purely data-driven approaches without appropriate domain knowledge incorporation risk missing important structure. The optimal path forward likely involves hybrid methodologies that combine domain-specific modeling assumptions with flexible machine learning components, leveraging complementary strengths rather than viewing these paradigms as competing alternatives.

In conclusion, deep learning approaches to stochastic volatility model calibration represent significant methodological advances offering clear practical benefits for computational efficiency and robustness, though they do not eliminate fundamental challenges inherent in inferring unobservable parameters from market prices. The careful characterization of numerical challenges in characteristic function evaluation and the detailed analysis of neural architectures and optimization algorithms provided in this work offer valuable guidance for researchers and practitioners implementing next-generation calibration systems. As the methodology matures and best practices emerge, neural network calibration seems likely to become a standard component of the quantitative analyst's toolkit, complementing rather than replacing traditional methods and enabling more sophisticated modeling with faster adaptation to evolving market conditions than previously possible.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Nafiu A, Balogun S O, Oko-Odion C, et al. Risk management strategies: Navigating volatility in complex financial market environments. World Journal of Advanced Research and Reviews, 2025, 25(1): 236-250.

[2] Ma Z, Chen X, Sun T, et al. Blockchain-based zero-trust supply chain security integrated with deep reinforcement learning for inventory optimization. Future Internet, 2024, 16(5): 163.

[3] Sun T, Yang J, Li J, et al. Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access, 2024.

[4] Patawari A, Das P. From traditional to computationally efficient scientific computing algorithms in option pricing: Current progresses with future directions. Archives of Computational Methods in Engineering, 2025: 1-40.
[5] Li Q, Jiang H, Lee C H. Phase-Space Generalized Brillouin Zone For Spatially Inhomogeneous Non-Hermitian Systems. Advanced Science, 2025: e08047.
[6] Zhang Y, Zhang H. Enhancing robot path planning through a twin-reinforced chimp optimization algorithm and evolutionary programming algorithm. IEEE Access, 2025, 12: 170057-170078.
[7] Gilli M, Maringer D, Schumann E. Numerical methods and optimization in finance. Academic Press, 2019.
[8] Devan M, Thirunavukkarasu K, Shanmugam L. Algorithmic trading strategies: real-time data analytics with machine learning. Journal of Knowledge Learning and Science Technology, 2023, 2(3): 522-546.
[9] Khalili H. Spatial competition of learning agents in agricultural procurement markets. Doctoral dissertation, Universitäts-und Landesbibliothek Bonn, 2019.
[10] Nikolenko S I. Synthetic data for deep learning. Cham, Switzerland: Springer, 2021, 174.
[11] Ge Y, Wang Y, Liu J, et al. GAN-Enhanced Implied Volatility Surface Reconstruction for Option Pricing Error Mitigation. IEEE Access, 2025.
[12] Hirsa A. Computational methods in finance. Chapman and Hall/CRC, 2024.
[13] Sabbar Y. A Review of Modern Stochastic Modeling: SDE/SPDE Numerics, Data-Driven Identification, and Generative Methods with Applications in Biology and Epidemiology. arXiv preprint arXiv:2508.11004, 2025.
[14] Shlezinger N, Whang J, Eldar Y C, et al. Model-based deep learning. Proceedings of the IEEE, 2023, 111(5): 465-499.
[15] Rohilla A. Exploring Volatility: Evolution, Advancements, Trends, and Applications. 2024.
[16] Alòs E, León J A. An intuitive introduction to fractional and rough volatilities. Mathematics, 2021, 9(9): 994.
[17] Bajaj P, Kaur J. Recent Developments and Applicability of the Black and Scholes Model in Option Pricing: A Literature Review. Mudra: Journal of Finance & Accounting, 2020, 7(2).
[18] Kristensen S. High-Performing Computational Methods for American Option Pricing and Implied Volatility Surface Construction. 2025.
[19] Cao W, Mai N T, Liu W. Adaptive knowledge assessment via symmetric hierarchical Bayesian neural networks with graph symmetry-aware concept dependencies. Symmetry, 2025, 17(8): 1332.
[20] Mai N T, Cao W, Liu W. Interpretable knowledge tracing via transformer-Bayesian hybrid networks: Learning temporal dependencies and causal structures in educational data. Applied Sciences, 2025, 15(17): 9605.
[21] Liu S, Borovykh A, Grzelak L A, et al. A neural network-based framework for financial model calibration. Journal of Mathematics in Industry, 2019, 9(1): 9.
[22] Chen S, Liu Y, Zhang Q, et al. Multi-Distance Spatial-Temporal Graph Neural Network for Anomaly Detection in Blockchain Transactions. Advanced Intelligent Systems, 2025: 2400898.
[23] Zhang Q, Chen S, Liu W. Balanced Knowledge Transfer in MTTL-ClinicalBERT: A Symmetrical Multi-Task Learning Framework for Clinical Text Classification. Symmetry, 2025, 17(6): 823.
[24] Ren S, Jin J, Niu G, et al. ARCS: Adaptive Reinforcement Learning Framework for Automated Cybersecurity Incident Response Strategy Optimization. Applied Sciences, 2025, 15(2): 951.
[25] Moradinia Z, Vandierendonck H, Murphy A. Navigating speed–accuracy trade-offs for multi-physics simulations. Meccanica, 2025, 60(6): 1583-1599.
[26] Mai N T, Cao W, Wang Y. The global belonging support framework: Enhancing equity and access for international graduate students. Journal of International Students, 2025, 15(9): 141-160.
[27] Liu Y, Ren S, Wang X, et al. Temporal logical attention network for log-based anomaly detection in distributed systems. Sensors, 2024, 24(24): 7949.
[28] Tan Y, Wu B, Cao J, et al. LLaMA-UTP: Knowledge-Guided Expert Mixture for Analyzing Uncertain Tax Positions. IEEE Access, 2025.
[29] Zheng W, Liu W. Symmetry-Aware Transformers for Asymmetric Causal Discovery in Financial Time Series. Symmetry, 2025, 17(10): 1591.
[30] Ivaşcu C F. Option pricing using machine learning. Expert Systems with Applications, 2021, 163: 113799.
[31] Hu X, Zhao X, Wang J, et al. Information-theoretic multi-scale geometric pre-training for enhanced molecular property prediction. PLoS One, 2025, 20(10): e0332640.
[32] Zhang H, Ge Y, Zhao X, et al. Hierarchical deep reinforcement learning for multi-objective integrated circuit physical layout optimization with congestion-aware reward shaping. IEEE Access, 2025.
[33] Wang J, Zhang H, Wu B, et al. Symmetry-Guided Electric Vehicles Energy Consumption Optimization Based on Driver Behavior and Environmental Factors: A Reinforcement Learning Approach. Symmetry, 17(6), 930.
[34] Hu, X., Zhao, X., & Liu, W. (2025). Hierarchical Sensing Framework for Polymer Degradation Monitoring: A Physics-Constrained Reinforcement Learning Framework for Programmable Material Discovery. Sensors, 2025, 25(14): 4479.
[35] Wang Y, Ding G, Zeng Z, et al. Causal-Aware Multimodal Transformer for Supply Chain Demand Forecasting: Integrating Text, Time Series, and Satellite Imagery. IEEE Access, 2025.
[36] Han X, Yang Y, Chen J, et al. Symmetry-Aware Credit Risk Modeling: A Deep Learning Framework Exploiting Financial Data Balance and Invariance. Symmetry, 2025, 17(3): 20738994.

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951
Online ISSN: 3007-696X
DOI: https://doi.org/10.61784/jtfe3062

INTEROPERABILITY TECHNOLOGIES AND CREDIT RECOGNITION MECHANISMS AMONG TRADE FINANCE BLOCKCHAIN NETWORKS

Yang Wu
HSBC Bank (China) Limited, Shanghai 200120, China.
Corresponding Email: wuyang2111@163.com

Abstract: Trade finance is the backbone of global commerce but remains technologically fragmented, leading to inefficiency, risk, and high costs. This study analyzes the technical and institutional barriers to interoperability among blockchain trade finance platforms, evaluates credit recognition mechanisms, and proposes a path toward seamless, secure financial integration. Comprehensive case studies from Asia and global consortia demonstrate that interoperability and automated credit recognition can reduce costs by 73–89%, settlement time from days to seconds, and significantly improve access for SMEs. Drawing on empirical data from eTradeConnect, mBridge, Marco Polo, and PBCTFP, we present a three-layer architectural framework integrating technical, institutional, and regulatory dimensions. The paper presents a phased implementation roadmap through 2028 and identifies policy recommendations for central banks, financial regulators, and industry consortia seeking to develop robust interoperability infrastructure.
Keywords: Blockchain interoperability; Trade finance; Credit recognition; Financial infrastructure

1 INTRODUCTION

Trade finance inefficiencies account for an estimated $2.5 trillion annual shortfall in global trade[1,2]. According to the World Bank, only 50% of all trade finance requests receive funding, with the remaining 50% going unfunded or redirected through informal channels[3]. The fragmentation between platforms, standards, and verification systems leads to duplicated due diligence, costly delays, and SME exclusion[4-6]. This inefficiency disproportionately impacts developing economies, where SMEs lack direct access to major financial institutions and struggle with outdated documentation processes.

Blockchains and distributed ledger technology (DLT) promise resolution through immutable audit trails, programmable settlement, and verifiable credit signals[7-9]. Early blockchain implementations in trade finance have attracted major banks and central banks to pilot platforms across Asia and globally. However, as adoption grows, interoperability between competing systems and robust, transparent credit recognition become critical[10-13]. The proliferation of incompatible blockchain platforms has created a "blockchain archipelago" rather than an integrated ecosystem—shippers in Manila cannot access credit facilities on the Hong Kong eTradeConnect network, central banks cannot settle payments using assets from competing networks, and importers must maintain separate accounts across multiple platforms.

This fragmentation defeats the original promise of blockchain technology: to eliminate intermediaries and streamline transactions. Instead, new intermediaries have emerged—multi-chain service providers, bridge operators, and liquidity aggregators—whose coordination failures risk undermining the benefits that DLT was designed to provide.

This paper addresses three core research questions: (1) What are the technical, institutional, and regulatory barriers to blockchain interoperability in trade finance? (2) How can standardized credit recognition mechanisms enable value transfer across heterogeneous blockchain networks? (3) What implementation pathways should central banks, regulators, and industry consortia pursue to build functional interoperability infrastructure?

2 LITERATURE REVIEW

2.1 Global Trade Finance Gaps and DLT Potential

The World Bank and BIS regularly report massive funding gaps and structural inefficiencies in international trade finance, especially for SMEs and developing economies[7,8]. According to the World Bank Trade Finance Program, an estimated $2.5 trillion financing gap exists annually, driven by three factors: (1) inefficient documentation processes requiring multiple intermediaries and creating delays of 7-10 days, (2) fragmented information systems where each institution maintains separate databases and requires customers to re-verify information, and (3) misaligned credit assessment frameworks where different jurisdictions apply different standards.

DLT offers distinct advantages addressing these gaps: immutable audit trails creating tamper-proof records, programmable settlement enabling simultaneous execution of multiple conditions, and verifiable credit signals based on transparent transaction history[7-9,14-17]. The transparency reduces fraud risk by 65% compared to paper- based systems, while cryptographic verification enables real-time validation without intermediaries. Early adopters like eTradeConnect have demonstrated measurable improvements in transaction throughput and cost efficiency[7], processing 563% more transactions year-over-year.

Central banks have recognized DLT's potential for cross-border payments. The Bank for International Settlements reports that CBDC-based settlement can reduce payment times from 2-3 business days to seconds, with 99.95% cost reduction[8,9]. This recognition has driven participation from 22+ central banks in initiatives like mBridge[15].

2.2 Interoperability Problems in Trade Finance Blockchain Adoption

Despite individual platform success, a critical problem has emerged: blockchain fragmentation. Studies highlight that parallel DLT initiatives—such as eTradeConnect, Marco Polo, mBridge, and PBCTFP—have rapidly proliferated[2,7,10,11,18]. This has led to data silos and technical incompatibility, increasing the need for robust cross-chain standards and data models[3,12,13]. Platform heterogeneity creates multiple coordination challenges[5,6].

Consensus Mechanism Incompatibility: Different blockchains use fundamentally different consensus mechanisms with different security assumptions and finality guarantees. Proof of Work uses computational puzzle-solving with probabilistic finality. Proof of Stake uses validator deposits with faster finality. Practical Byzantine Fault Tolerance uses voting with immediate finality. Proof of Authority uses designated validators. These mechanisms require reconciliation in any interoperability protocol, either accepting weaker finality guarantees or adding mechanisms strengthening weak-finality models.

Smart Contract Language Incompatibility: Different blockchains use different execution models and programming languages. Solidity and Ethereum use state mutations with reentrancy risks. Hyperledger Fabric uses Go/Java with different execution models. Cardano uses functional programming paradigms. Corda uses deterministic contracts. Smart contract logic must be rewritten for each platform, creating maintenance burden and security risks.

Data Schema Incompatibility: Different platforms represent trade finance concepts differently. eTradeConnect uses UN/CEFACT standards. Marco Polo uses proprietary Corda models. we.trade aligns with SWIFT standards. mBridge uses CBDC-specific schemas. Semantic mapping between these schemas is non-trivial, requiring careful data modeling to identify corresponding fields across systems.

Finality and Settlement Certainty: Trade finance requires absolute settlement finality. Probabilistic finality blockchains (Bitcoin, Ethereum before sharding) cannot be used directly for irreversible settlements. Deterministic finality blockchains (Hyperledger Fabric, Corda) provide immediate finality. Interoperability protocols must map weaker-finality transactions to stronger-finality settlement mechanisms.

2.3 Credit Recognition Mechanisms

Traditional, centralized credit rating systems introduce delay and bias[1,19]. Credit rating agencies (Moody's, S&P Global, Fitch) assign ratings based on historical financial data, industry analysis, and proprietary algorithms. Banks maintain detailed creditworthiness information accumulated through customer relationships. This centralized system functions reasonably for large established corporations but creates significant problems for SMEs: approximately 95 million SMEs lack access to formal credit due to insufficient credit history, collateral, or banking relationships. Emerging market firms with equivalent creditworthiness to developed market firms often face significantly higher financing costs.

Ratings update quarterly or annually, missing rapid creditworthiness changes. During economic downturns, agencies tend to downgrade simultaneously, reducing credit availability when most needed.

Blockchain experimentation has led to proof-of-reputation, DeFi collateral, and KYC relay solutions[9,14,16,20]. On chain transaction history creates immutable cryptographic records where a participant with 10,000 completed transactions worth $500 million with zero defaults demonstrates creditworthiness through verifiable history rather than third-party assessment. This approach provides transparency (all participants can verify history), immutability (history cannot be altered retroactively), real-time updates (reputation changes immediately), and standardization (all assessed through identical mechanisms).

However, practical integration with AML/CFT and Basel frameworks is nascent[17,21,22]. Automated regulatory reporting and on-chain verification offer promise for reducing compliance burden[20,22], but standardized procedures for cross-chain credit information sharing remain underdeveloped.

2.4 Empirical Studies and Gaps

Field studies document measured gains from DLT adoption. eTradeConnect reduced documentation fraud by >70% and settlement costs by 73%[7]. mBridge reduced cross-border settlement from 2-3 business days to 3.7 seconds[8,9]. Marco Polo's multi-party smart contracts reduced documentary credit processing from 7-10 days to 2.3 days[10,11].

However, multi-platform coexistence and integration remain under-addressed in the literature[3-6]. Most studies examine individual platforms in isolation rather than analyzing cross-chain coordination.

3 THEORETICAL FRAMEWORK

3.1 Defining Interoperability

Drawing on UN/CEFACT and ISO 20022 standards[12,13], we define interoperability as the seamless, standards-based exchange of value and credit data between separate systems, with verifiable finality, regulatory equivalence, and commercial enforceability[10,11]. This definition encompasses three distinct dimensions:

Technical Interoperability: The ability to transmit data and execute transactions across distinct blockchain networks. Requires standardized communication protocols enabling messages from one network to be verified on another, asset wrapping mechanisms enabling value transfer, standardized data schemas enabling consistent information interpretation, and cross-chain execution protocols.

Institutional Interoperability: The organizational and governance frameworks enabling multiple independent blockchain networks to coordinate operations. Requires standardized operational procedures and settlement protocols, shared governance mechanisms or coordination forums, aligned regulatory compliance procedures, and interoperable audit and reporting systems.

Semantic Interoperability: The capacity of different blockchain systems to consistently interpret transaction data, asset valuations, and credit assessments. Requires standardized asset definitions and valuation frameworks, consistent ontologies for trade finance concepts, aligned temporal reference points and settlement finality definitions, and common credit scoring methodologies. Effective interoperability requires alignment across these three dimensions plus integration with regulatory compliance frameworks[18,21,22].

3.2 Three-Layer Architectural Model

The paper adopts a comprehensive three-layer model addressing the full spectrum of interoperability requirements:

Layer 1 - Technical Interoperability: Standardized protocols, data structures, and cross-chain messaging mechanisms[12,13,16]. This layer implements standardized cross-chain messaging enabling reliable transaction initiation and settlement verification across networks. Protocol operation follows: (1) Transaction Initiation—participant initiates transaction on Network A specifying source details, destination details, settlement terms, amount, and timestamp; (2) Cross-Chain Relay—authorized relay nodes observe transaction and create cryptographic proof confirming occurrence; (3) Verification and Finality—target network verifies relay signature and cryptographic proof; upon verification, transaction achieves settlement finality and cannot be reversed; (4) Settlement Execution—target network executes settlement action and confirms completion, with both networks recording transaction completion in audit logs.

Layer 2 - Institutional Coordination: Consortium governance, credit information sharing, credit relay protocols, and collateral management[2,21]. Rather than each institution independently verifying all customers, participating institutions form verification consortia. Each institution verifies customers within its geographic jurisdiction or sector expertise. The consortium maintains shared registry containing verified participant information including participant identifier, legal name, jurisdiction, verifying institution, verification date, beneficial ownership data, sanctions status, and credit profile (transaction volume, completion rates, default history, average payment days, current credit rating).

Layer 3 - Regulatory and Compliance: KYC/AML compliance, Basel III capital requirements, and policy alignment with IMF/World Bank frameworks[1,2,18,21,22]. This layer ensures credit recognition mechanisms comply with financial regulations and maintain systemic stability. Smart contracts integrate with regulatory reporting systems enabling real-time transaction reporting to central banks and regulators, aggregate position reporting for systemic risk monitoring, automatic sanctions screening with transaction blocking for sanctioned participants, and capital requirement calculations based on counterparty credit risk.

3.3 Mechanisms of Credit Recognition in Decentralized Systems

Theoretical mechanisms for decentralized credit recognition include:

On-Chain Transaction History: Every transaction creates immutable cryptographic records. A participant with 10,000 completed transactions worth $500 million with zero defaults demonstrates creditworthiness through verifiable history rather than third-party assessment[16,17].

Consensus-Based Performance Verification: Smart contracts create verifiable records of condition satisfaction. A shipper consistently meeting delivery timelines encoded in contract conditions, an importer consistently paying invoices before due dates, or a logistics operator consistently providing accurate information each creates verifiable performance evidence.

Cryptographic Collateral Mechanisms: Rather than assessing creditworthiness through historical records, blockchain enables lending based on collateral. A participant locks $150 of cryptocurrency as collateral, borrows $100, and if they default, the collateral is automatically liquidated to repay the loan[3,6]. Over-collateralization eliminates credit risk, and as borrowers demonstrate reliable repayment, collateralization requirements decrease progressively.

Network-Based Reputation Through Relationship Verification: Rather than assessing individual credit independently, systems assess creditworthiness through verified network relationships. If Bank A has completed $1 billion in transactions with Bank B over 10 years with zero defaults, other participants can accept Bank B's creditworthiness claims based on this network relationship[19].

All mechanisms must be mapped to regulatory constraints[21,22] to ensure Basel III compliance and AML/CFT adherence.

4 DATA AND METHODOLOGY

4.1 Data Sources

This study draws on multiple data sources:
Case study metrics from official institutions: HKMA (eTradeConnect)[7], BIS & HKMA (mBridge)[8,9], R3/TradeIX (Marco Polo)[10,11], BOCHK (PBCTFP)[14].
Platform maintenance documentation: Workflow diagrams, transaction data, and publicly available field reports[[3-5,7-10].
Regulatory and policy documents: From Basel Committee[21], IMF[2], World Bank[1], European Commission[18], and Asian financial authorities[15-17].
Academic and industry research: Leading consultancies including Deloitte[3], McKinsey[4], Boston Consulting Group[5], and Goldman Sachs[6].

4.2 Empirical Strategy

Our empirical approach employs multiple complementary methods:
Comparative Platform Analysis: We analyze features, throughput, cost/time savings, and fraud reduction across eTradeConnect, Marco Polo, mBridge, and PBCTFP[7-11,14]. Metrics include transactions per second (TPS), average settlement time, fraud incident rates, documentation cost per transaction, and participant satisfaction scores.
Cross-Case Synthesis: We examine KYC/AML procedures, settlement mechanisms, and credit relay protocols implemented across platforms[7-10]. This approach identifies common patterns and platform-specific innovations.
Statistical Compilation: Transaction volume, settlement time, fraud rates, and cost metrics are compiled from platform reports and published announcements[7-9,20,23]. We normalize metrics across platforms to enable comparison despite different operational models.
Normative Regulatory Comparison: Assessment against Basel III[21], MiCA[18], GDPR[17], and international standards[12,13] identifies regulatory gaps and opportunities.
Cost-Benefit Analysis: Quantification of efficiency gains and risk reduction from interoperable platforms relative to traditional trade finance[1-4]. We calculate net present value of interoperability investments over 5-year and 10-year horizons.

5 EMPIRICAL RESULTS

5.1 Major Blockchain Trade Finance Platforms: Technical Attributes

Table 1 Key Technical Attributes of Leading Platforms

Platform	Consensus	TPS	Settlement Time	Finality	Participants	Launch
eTradeConnect	PBFT (Fabric)	3,500	3.2 days	Strong	39	2018
Marco Polo	PBFT (Corda)	500	2.3 days	Strong	200+	2017
mBridge	CB Validators	High	3.7 seconds	Absolute	22+	2024
PBCTFP	Fabric	2,000	2.8 days	Strong	17	2020

The table 1 reveals key differences: eTradeConnect and PBCTFP, both using Hyperledger Fabric, demonstrate high throughput (2,000-3,500 TPS) but settlement times in the 3-day range reflecting traditional banking integrations. Marco Polo's Corda implementation achieves lower throughput (500 TPS) but focuses on legal certainty through deterministic finality. mBridge's central bank validators enable absolute finality with unprecedented settlement speed (3.7 seconds), though throughput metrics are still evolving as the system scales.

5.2 Platform Interoperability and Settlement Results

eTradeConnect-PBCTFP Pilot Results
Interoperability pilots between eTradeConnect and PBCTFP reduced end-to-end settlement from 12 days to under 3 days, with documentation fraud costs falling from $450 to $120 per transaction—a 73% reduction[7,14]. Real-time KYC/AML verification through shared registries eliminated duplicate compliance checking. The pilot processed 1,247 transactions totaling $1.8 billion in 2024, with zero settlement failures and <0.01% fraud rates[7].
mBridge Cross-Border Settlement
mBridge, connecting 22 central banks, achieved 99.95% settlement time reduction, with average settlement accelerating from 2-3 business days to 3.7 seconds[8,9]. Transaction volume reached $14.2 billion in 2024[8]. Cross-border payment costs fell from $50-100 per transaction to $0.50-1.00[9]. The system demonstrated 99.97% transaction success rate with

<0.01% failed settlements[9]. Central bank governors have confirmed mBridge's capability for medium-term production deployment, with expansion to additional jurisdictions planned.

Marco Polo Multi-Party Coordination

Marco Polo's multi-party smart contracts reduced documentary credit processing from 7-10 days to 2.3 days, with 200+ participating banks confirming operational efficiency[10,11]. The platform has processed cumulative volumes exceeding $1.2 trillion in committed credit lines. Commercial pilots involving major corporations like Voith and KSB demonstrated that payment commitments could be secured through digital data exchange matching previously agreed data, triggering automatic payment obligations[10].

5.3 Impact on Credit Recognition and SME Inclusion

Implementing on-chain verified creditworthiness based on transaction history reduced SME rejection rates by 45%—from 50% historical baseline to 27.5% with blockchain verification[2,19]. This means approximately 42.75 million SMEs gained access to trade finance who previously faced rejection. Participating SMEs obtained more favorable financing terms as their on-chain reputation scores improved, with borrowing costs declining by average 180 basis points after 18 months of transaction history[1].

Collateralization mechanisms, particularly in mBridge and PBCTFP, enabled credit extension to SMEs with limited prior borrowing records[9,14,19]. By locking cryptocurrency or tokenized assets as collateral, SMEs could access working capital without extensive credit history. As participants demonstrated reliable repayment, collateralization requirements decreased—many SMEs reduced collateral ratios from 150% to 110% after 12 months of zero-default history[19]. Automated reputation scoring reduced credit information asymmetry by 73% compared to traditional third-party rating assessments[19]. Participants with verified transaction history could demonstrate creditworthiness through cryptographic proof rather than relying on rating agency assessment. This particularly benefited emerging market participants systematically underrated by traditional agencies—on-chain credit scores averaged 2.3 notches higher than Moody's equivalents for comparable emerging market firms[19].

5.4 Compliance and Risk Management Achievements

All interoperable platforms achieved instant, compliance-logged KYC and AML screening against sanctions lists[7-9][17,20]. Transactions involving sanctioned parties were automatically blocked, with <0.5% false-positive rate requiring manual review[7,9]. Sanctions list updates from regulatory authorities were ingested and applied in real-time.

Automated regulatory reporting reduced post-trade compliance labor by 61%[12,23]. Rather than batch processing requiring days or weeks, transactions generated compliance reports immediately upon settlement. Regulators gained real-time visibility into transaction flows; suspicious transactions were identified immediately rather than through periodic batch analysis[22].

GDPR compliance was achieved through off-chain personal data storage with on-chain cryptographic identifiers, satisfying European data localization requirements while maintaining blockchain immutability[13]. The "right to be forgotten" was implemented through off-chain data deletion with blockchain records persisting, achieving regulatory compliance while maintaining transaction history.

Basel III capital requirements were updated to reflect blockchain-based credit verification[21]. Banks could reduce reserve requirements by 15-20% for counterparties with verified on-chain creditworthiness and zero-default history, compared to unverified counterparties requiring 100% reserves. This enabled banks to deploy capital more efficiently toward productive lending rather than maintaining idle reserves[13].

6 ANALYSIS OF INTEROPERABILITY MECHANISMS

6.1 Eight Approaches to Cross-Chain Coordination

The technical literature identifies eight distinct interoperability mechanisms:

Relay Chains: Maintain verified block headers from source blockchains, enabling verification of transactions without requiring complete blockchain history downloads[1,22]. Advantages: no modifications required to source blockchains, strong cryptographic verification. Disadvantages: relay chains must process headers from all source chains (scalability constraint), latency equals source chain block time plus relay processing.

Sidechains and Pegged Assets: Maintain fixed exchange rates with parent blockchains through peg mechanisms. Assets locked on parent chains result in equivalent pegged asset issuance on sidechains[13]. Advantages: enables high throughput with security anchoring, allows different consensus mechanisms. Disadvantages: introduces custodial risk, requires honest operator majority, creates valuation differences.

Atomic Swaps: Enable cross-chain transactions where both transactions complete or both rollback, eliminating counterparty risk. Time-locked conditions coordinate parties: Party A creates cryptographic commitment, both create reciprocal locking conditions, Party A reveals secret on target chain, Party B uses secret on source chain[12].

Notary Schemes: Use trusted validator sets observing multiple blockchains and attesting to transaction occurrence. Validators independently verify occurrence, create signatures, submit M-of-N multisignatures to target blockchain[7,10].

Liquidity Pools: Maintain asset reserves enabling exchanges based on predefined pricing formulas[9,16]. Advantages: efficient atomic exchanges, liquidity provider incentives. Disadvantages: requires capital lockup, creates impermanent loss risk.
Wrapped Assets: Represent claims on assets held in regulated custodial accounts. Institutions deposit assets in escrow, mint equivalent wrapped tokens on target blockchains, enable redemption through custody withdrawal[14].
Cross-Chain Oracles: Relay information from external sources into smart contracts. Oracle nodes observe external data, create cryptographically-signed attestations, submit to target blockchains[16,17].
Standardized Bridging Protocols: Standardize cross-chain messaging, combining elements of above approaches. Bridges follow: observation phase (validators monitor source blockchain), verification phase (validators verify transaction authenticity), attestation phase (create signatures), transmission phase (aggregate and transmit), verification phase (target blockchain verifies), execution phase (authorized action executes)[7-11].

6.2 Institutional Coordination Models

Successful interoperability requires institutional coordination beyond technical protocols.
Consortium-Based Governance: Rather than centralized control, participating institutions form consortia where decisions are made through consensus or qualified majority voting[7-9]. eTradeConnect is governed by HKMA coordinating 39 participating banks. mBridge is governed by participating central banks. Marco Polo is governed by member banks through R3. This distributed governance prevents single-institution dominance while enabling rapid decision-making compared to regulatory processes.
Shared Credential Infrastructure: W3C Decentralized Identifiers (DIDs) enable portable digital identities not bound to single institutions[16,17]. A participant can maintain a DID aggregating credit information from multiple sources. W3C Verifiable Credentials allow cryptographic proof of creditworthiness to be issued by trusted institutions [17]. For example, a bank can issue a credential stating, "Participant ABC has completed $50 million in transactions with zero defaults over a two-year period," which is cryptographically signed by the issuing bank. Other institutions can verify the signature and accept the credential without needing to access the underlying transaction details.
Verification Consortia: Rather than each institution independently verifying all customers, institutions form consortia where each verifies customers within its jurisdiction[7,17]. Bank of Thailand verifies Thai importers and exporters. Bank Negara Malaysia verifies Malaysian businesses. HKMA verifies Hong Kong entities. This approach provides efficiency (each institution verifies only local customers), regulatory alignment (verification follows local standards), privacy protection (institutions share credential summaries rather than raw data), and real-time updates (changes propagate immediately)[17].

7 THREE-LAYER FRAMEWORK IMPLEMENTATION

7.1 Technical Layer: Standardized Messaging

ISO 20022 provides foundation for technical interoperability[12,13]. The standard defines message formats for cross-border credit transfers (pacs.009), payment status reporting (pacs.028), and transaction inquiries (camt.027). Participating platforms should implement ISO 20022-compatible messaging enabling direct data exchange without transformation layers. Cross-chain message passing protocol operates through: (1) Transaction Initiation on source blockchain specifying recipient, amount, conditions, and settlement terms; (2) Relay observation and cryptographic proof creation; (3) Target blockchain verification against known validator signatures; (4) Atomic settlement where transaction either fully completes or fully reverses, eliminating settlement risk.

7.2 Institutional Layer: Shared Registries and Governance

KYC/AML credential sharing through consortium registries eliminates duplicate verification. When a participant moves between jurisdictions, rather than requiring full re-verification, receiving institution can accept credential issued by verified counterpart, reducing onboarding time from weeks to days[7,17].
Collateral valuation standards enable cross-chain collateral recognition. A participant's cryptocurrency holdings, tokenized assets, or traditional collateral can be recognized across platforms when valued according to standardized frameworks with real-time price feeds from multiple independent sources[9,14].
Settlement priority sequencing ensures orderly settlement when multiple transactions compete for limited liquidity. Smart contracts implement priority based on payment instructions, settlement currency preferences, and counterparty creditworthiness[7,10].

7.3 Regulatory Layer: Compliance Automation

Automated sanctions screening blocks transactions involving sanctioned parties. Regulatory authorities provide sanctions lists that participating platforms ingest and apply in real-time[7-9,17,20].
Capital requirement calculations incorporate counterparty credit risk using on-chain verification data[21]. Basel III framework updated to recognize blockchain-based credit verification, enabling reduced capital reserves for verified counterparties[21].

Regulatory harmonization mechanisms map conflicting requirements across jurisdictions. When Hong Kong requires specific data retention policies and EU requires different policies, system applies Hong Kong policies to Hong Kong participants and EU policies to EU participants[17,18].

8 IMPLEMENTATION ROADMAP (2025-2028+)

8.1 Phase 1: Foundation (2025-2026)

Objectives: Establish interoperability standards, build bridging infrastructure, develop regulatory frameworks.
Key Activities:
Standards Development (2025): ISO technical committee publishes ISO 20022-compatible interoperability standards[12,13]. UN/CEFACT develops blockchain semantics framework for trade finance[13]. SWIFT publishes gpi+ protocol enabling CBDC settlement[15].
Infrastructure Deployment (2025-2026): Central Bank Digital Currency networks go live in 15+ countries[19]. eTradeConnect expands to 8 additional countries. mBridge connects 22 central banks[8,9]. Marco Polo adds 100+ new member banks[10,11].
Regulatory Framework (2025-2026): G7 publishes blockchain financial regulation guidelines. Basel Committee issues guidance on blockchain credit risk assessment[21]. EU finalizes MiCA implementation[18]. Asian regulators establish mutual recognition agreements.
Expected Outcomes: 5 major trade finance blockchains achieve basic interoperability. 50+ participating central banks on CBDC networks. $50 billion in cross-chain settlement volume.

8.2 Phase 2: Integration (2026-2027)

Objectives: Connect major platforms through standardized bridges, develop cross-chain credit recognition, achieve mainstream financial institution adoption.
Platform Integration: Deploy standardized bridging protocols connecting eTradeConnect, PBCTFP, Marco Polo, mBridge, and emerging platforms[7-11,14]. Implement standardized KYC/AML sharing infrastructure[21]. Enable direct cross-chain asset transfers[12,13].
Credit Recognition: Establish decentralized credit scoring using on-chain transaction history[19]. Deploy standardized collateral valuation mechanisms. Implement cross-chain collateral recognition[9,14].
Adoption Expansion: Regional development banks integrate with trade finance blockchains. 500+ additional financial institutions adopt blockchain platforms[3,4]. 1,000+ SMEs access blockchain-based trade finance[1,19].
Expected Outcomes: $500 billion annual cross-chain settlement volume. 80% cost reduction in trade finance. 85% settlement time reduction (7-10 days to 12-18 hours).

8.3 Phase 3: Optimization (2027-2028)

Objectives: Achieve full interoperability, automate settlement and credit recognition, establish regulatory parity.
Interoperability Enhancement: Implement atomic cross-chain settlement (simultaneous completion or rollback)[12,13]. Deploy cryptographic verification systems for privacy-preserving verification[16,17]. Enable smart contract execution coordination across multiple blockchains[10,11,14].
Automation: 95%+ of trade finance transactions execute with zero human intervention[7-9]. Automatic collateral rebalancing across platforms[9,14]. Real-time credit score updates[19].
Regulatory Harmonization: IMF/World Bank establish unified blockchain financial regulation framework[1,2]. Mutual regulatory recognition agreements among major jurisdictions[17,18,21,22]. Standardized AML/CFT compliance across platforms[17,20].
Expected Outcomes: $5+ trillion annual blockchain-based trade finance. Near-elimination of trade finance documentation delays. Full cost/time parity between blockchain and traditional settlement.

8.4 Phase 4: Evolution (2028+)

Objectives: Blockchain becomes default settlement mechanism, enable new financial services, manage systemic risk across networks.
Service Innovation: Real-time trade finance (settlement within hours). Programmable supply chain finance (automated financing at each supply chain stage). Synthetic asset trading (direct trading of trade finance instruments)[1,9,10].
Systemic Risk Management: Cross-chain stress testing simulating cascading failures. Systemic risk monitoring across connected networks enabling early intervention. Emergency intervention procedures for major disruptions[7-10].

9 DISCUSSION AND IMPLICATIONS

9.1 Technical Interoperability Achievement

The empirical results demonstrate that technical interoperability using standardized protocols (ISO 20022, UN/CEFACT standards) is operationally feasible[12,13]. Cross-chain message passing, relay mechanisms, and notary schemes enable reliable settlement finality across heterogeneous platforms[7-11]. The eTradeConnect-PBCTFP pilot and mBridge implementation confirm that interoperability can be achieved in production environments.

However, consensus mechanism incompatibility and smart contract language fragmentation remain coordination challenges[3-5]. Developing language-agnostic smart contract standards or transpilers would reduce maintenance burden and security risks. Standards bodies should prioritize this work to enable more seamless cross-chain integration.

9.2 Institutional and Regulatory Harmonization

Consortium-based governance models (HKMA coordination for eTradeConnect, BIS facilitation for mBridge) have proven effective in coordinating institutional participation[7-9]. These models enable rapid decision-making while maintaining institutional autonomy. However, expanding to include non-bank institutions, payment service providers, and fintech requires governance model evolution.

Regulatory harmonization through mutual recognition agreements between jurisdictions is critical but incomplete[17,18,21,22]. The EU's MiCA regulation provides a template that other jurisdictions should consider adopting[18]. However, fundamental conflicts remain—GDPR data localization requirements conflict with blockchain immutability, requiring creative solutions like off-chain data storage with on-chain identifiers.

9.3 Credit Recognition and Financial Inclusion

Decentralized credit recognition mechanisms demonstrate measurable benefits for SME inclusion and cost reduction[1,19]. On-chain transaction history provides objective, real-time creditworthiness assessment without bias from traditional rating agencies[19]. The 45% reduction in SME rejection rates[1,19] represents significant progress toward financial inclusion, particularly benefiting developing economy SMEs historically underserved by traditional rating agencies.

However, bootstrapping initial credit scores for new market entrants remains challenging and requires transition mechanisms. Hybrid approaches combining traditional ratings with on-chain verification could provide gradual transition during early adoption phases. Over time, as transaction volumes accumulate, on-chain assessment could gradually replace traditional ratings.

10 CONCLUSION

Interoperability and reliable credit recognition are the pivotal next steps for global trade finance DLT advancement[7,8,10,11]. Platform pilots confirm major gains in efficiency, risk controls, and SME inclusion[1,7-9,14]. The three-layer framework (technical, institutional, regulatory) points to a road map that can yield scalable benefits, provided regulatory harmonization and neutral technical standards are adopted industry-wide[12,13,18].

The evidence suggests that by 2028, interoperable blockchain platforms could consolidate into 3-5 major networks, processing $5+ trillion in annual trade volume with near-universal cost-efficiency and minimal fraud[1-4].

Settlement times could approach real-time, while fraud rates remain below 0.01%. SME access to trade finance could expand by 45%+ through decentralized credit recognition enabling previously excluded participants to demonstrate creditworthiness.

Policy Recommendations:

1. **Adopt ISO 20022 and UN/CEFACT standards** as mandatory frameworks for blockchain trade finance platforms[12] [13]. Regulatory bodies should require compliance and establish testing procedures.
2. **Establish international interoperability working groups** convening central banks, regulators, and industry representatives to develop harmonized frameworks and resolve jurisdictional conflicts[17,18,21,22].
3. **Update Basel III guidance** to formally recognize blockchain-based credit verification, enabling banks to reduce capital reserves for verified counterparties and deploy capital more efficiently[21].
4. **Develop mutual regulatory recognition agreements** among major jurisdictions enabling participants verified in one jurisdiction to operate in others without duplicative verification[17,18].
5. **Invest in research** on post-quantum cryptography, consensus mechanism reconciliation, and smart contract language standardization to address emerging technical challenges[3-5,12].

The path toward fully interoperable, credit-recognized trade finance infrastructure is technically achievable, institutionally viable, and economically compelling. Decisions made now about interoperability standards, credit recognition mechanisms, and regulatory frameworks will structure international trade finance for decades to come. The industry faces a critical window—within 2-3 years, early incompatible implementations will become difficult to change. Standardization now prevents costly future migration challenges.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] World Bank. Trade Finance Program: Annual Report 2024. Washington, D.C., 2024.
[2] International Monetary Fund. Blockchain Technology in Global Financial Infrastructure. IMF Working Papers, 2024.
[3] Deloitte. Global Blockchain in Banking Survey. Deloitte Global Report, 2024.
[4] McKinsey & Company. Blockchain Adoption in Global Trade Finance. McKinsey & Company Report, 2024.
[5] Boston Consulting Group. Future of Trade Finance Infrastructure. BCG Report, 2024.
[6] Goldman Sachs. Blockchain Technology Impact on Cross-Border Payments. Goldman Sachs Report, 2024.
[7] Hong Kong Monetary Authority. The launch of eTradeConnect and the Collaboration with we.trade. HKMA Press Release, October 31, 2018.
[8] Hong Kong Monetary Authority. Project mBridge reaches MVP stage. HKMA Press Release, June 5, 2024.
[9] Bank for International Settlements. Project mBridge reached minimum viable product stage. BIS Innovation Hub, August 2024.
[10] R3/TradeIX. Marco Polo trade finance network sees successful commercial pilot. Commercial Payments International, July 8, 2020.
[11] R3 Consortium. Marco Polo: Leading Global Banks Together With TradeIX and R3. Press Release, February 21, 2018.
[12] UN/CEFACT. Blockchain Semantics for International Trade. United Nations Economic Commission for Europe, 2024.
[13] ISO. ISO 20022: Financial Industry Business Messages Standards. International Organization for Standardization, 2024.
[14] China Industrial Bank. mBridge Project - Multi-Central Bank CBDC Bridge. CIB News, 2024.
[15] Bank of Thailand. mBridge: Multi-Central Bank CBDC Network. Bank of Thailand Publications, 2024.
[16] Monetary Authority of Singapore. Financial Infrastructure Modernization: Blockchain Integration Strategy. MAS Publications, Singapore, 2024.
[17] Bank Negara Malaysia. Trade Finance Blockchain Integration Report. BNM Publications, Kuala Lumpur, 2024.
[18] European Commission. MiCA Regulation: Digital Finance Framework. EU Technical Standards, 2024.
[19] ICC Banking Commission. Uniform Customs and Practice (UCP 600) Blockchain Implementation. ICC Publications, Paris, 2024.
[20] Reuters. Trade Finance Goes Digital: Blockchain Adoption Accelerates. Reuters News, 2024.
[21] Basel Committee on Banking Supervision. Basel III Blockchain Credit Risk Guidance. BIS Publications, November 2024.
[22] European Banking Authority. Draft technical standards on prudential treatment of crypto asset exposures. EBA Publications, January 2025.
[23] Financial Times. Central Banks Race to Develop CBDC Interoperability. Financial Times News, 2024.

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951
Online ISSN: 3007-696X
DOI: https://doi.org/10.61784/jtfe3063

FORECASTING STOCK PRICES WITH DEEP LEARNING MODELS: A COMPARISON OF LONG-SHORT TERM MEMORY (LSTM), GATED RECURRENT UNIT (GRU), ATTENTION MECHANISM, AND TRANSFORMER MODE

ZeTong Li[1], JiuRu Lyu[2], ZiHan Wang[3], Liu Yang[3*]
[1]Department of Electronic Science and Technology, Xi'an Jiaotong-Liverloop University, Suzhou 215000, Jiangsu, China.
[2]Department of Mathematics, Emory College of Art and Science Emory University, Atlanta, United States.
[3]School of Mathematics and Physics, Xi'an Jiaotong-Liverloop University, Suzhou 215000, Jiangsu, China.
Corresponding Author: Liu Yang, Email: liu.y73612@gmail.com

Abstract: With the expansion of the stock market, more and more people have started to use deep learning models to predict the stock market and facilitate their trading decisions. This paper compares four mainstream deep learning models for stock price prediction: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Attention Mechanism, and Transformer Model. Using MSE and RMSE as the evaluation metrics, we found LSTM performs the best in stock price prediction of the four companies of selection: Boeing Co, General Electric Co, Coca-Cola Co, and Johnson & Johnson. With a deeper analysis of the result, we found several limitations of LSTM, such as inconsistency of accuracy when forecasting the stock price of different firms. Hence, we suggested corresponding ways of improvement: adding more training data, introducing external factors, and integrating LSTM with other models.
Keywords: Deep learning; LSTM; GRU; Attention; Transformer; Stock; Forecasting; BA; GE; KO; JNJ

1 INTRODUCTION

With the stock market's expansion, more people are involved in stock trading. In 2019, about 600 million people worldwide bought stocks, and global stock transactions reached 60.36 trillion US dollars. Whether the stock market can be predicted has attracted more and more attention because effective prediction of the volatility of stock prices can not only strengthen financial risk management but also increase investors' enthusiasm in decision-making. In recent years, artificial intelligence, such as deep learning technology, has been combined with the financial industry to build some models to predict and analyze the stock market's volatility, improving the accuracy of stock price volatility prediction. However, because the financial market is a nonlinear and ever-changing complex dynamic system, we still can not accurately predict the changes in the stock market.

In the past, economists have been devoted to explaining economic phenomena, hoping to find the laws of economic operation existing in economic phenomena. Therefore, the previous methods of predicting stock prices, which usually need professional financial knowledge and only an understanding of uncomplicated time sequences in finance, need help to make perfect predictions [1]. Instead, deep learning technology has apparent benefits in dealing with complex and ever-changing problems, so more researchers intend to use it to make predictions. Recently, some models such as LSTM, Attention, and Transformer have been designed to predict stock prices and made some achievements. In 2017, Nelson et al. first used the LSTM model to make predictions [2]. However, following their work, in 2019, Li et al. found that LSTM could not obtain long-term dependence in long-term time series because it is limited by distinguishable position [3]. In 2019, Qiu et al. used wavelet transform to process stock data and LSTM neural network based on attention to forecasting the opening price of stocks and achieved good results. In addition, some researchers have also improved the LSTM model [4]. Li et al. 2018 proposed a multi-input LSTM element that can differentiate the mainstream factors from the auxiliary ones and perform better than the traditional LSTM model [3]. Furthermore, in 2017, Vaswani et al. created a sequence-to-sequence model called 'transformer' that adopts a multi-head self-attention mechanism to improve its ability to learn long-term dependence [5]. Ding et al. 2020 improved the ability of the original Transformer model to seize the short-term, long-term, and hierarchical dependence of financial time sequence [1].

With the continuous improvement of basic models such as LSTM, more and more advanced models have been developed and applied to stock price forecasting. However, each model has advantages and disadvantages, so finding the most suitable model is essential. In this paper, we used LSTM, GRU, Attention, and Transformer models to predict the stocks of General Electric Company, Johnson & Johnson, Coca-Cola, and Boeing. We calculated the values of the mean square error (MSE) and root mean square error (RMSE) to find out which model is the most accurate.

The rest of our paper is organized as follows: In Section II, we collect the literature related to this field. Section III walks through the related theories behind each model. Section IV introduces how to collect data, conduct experiments and analyze the experimental results. In section V, We draw our experimental conclusions and our ideas for future research fields.

2 LITERATURE REVIEW

Long Short-Term Memory (LSTM), a recursive network structure with an appropriate gradient-based learning algorithm, was first proposed by Hochreiter and Schmidhuber [6]. The advantage of LSTM is that it can handle noises, distributed representations, and continuous values [6]. Nevertheless, it also has limitations, such as the difficulty of solving problems like the strongly delayed XOR problem [3].

Several alternative models were proposed to improve the limitations of LSTM. For example, Li et al. proposed a new MI-LSTM model that enabled the mainstream to determine the use of other factors and to use a dual-stage attention mechanism for hidden states with different memory cell inputs and different time steps to improve accuracy [3].

Continuing to improve the performance of deep learning models to predict the stock market, Gupta proposed a new data expansion method in the GRU-based StockNet model, which consists of an injection module to prohibit overfitting and a survey module for stock index forecasting [7]. Compared with other models, this model has significantly lower RMSE, MAE, and MAPE [7]. Meanwhile, the GRU-based in-network data augmentation method is the unique feature of this study [7].

Apart from LSTM and GRU, the Transformer model is also a mainstream model for stock price prediction. Wang et al. proved that the Transformer model is better than traditional deep learning models in forecasting accuracy and net worth analysis [8]. Because the Transformer has a more vital ability to collect critical features and gets better prediction performance, Wang et al. inferred that financial time series prediction is a promising application field of transformer architecture [8]. In practice, by predicting transformers, investors can obtain higher excess returns [8].

To improve on the Transformer model, Ding et al. proposed some improvements to the Transformer model, including Multi-Scale Gaussian Prior, Orthogonal Regularization, and Trading Gap Splitter [1]. Their proposed Transformer-based method is superior to several advanced baselines in two fundamental trading markets compared with models such as CNN, LSTM, and ALSTM [1].

Inspired by the attention mechanism in biological phenomena, studies also reveal that attention mechanisms can be successfully integrated with other deep learning models. For instance, Zhang and Zhang proposed to optimize the LSTM model using the attention mechanism to improve its accuracy in predicting stocks [9]. The three models were also evaluated using K-fold cross-validation, and the LSTM-Attention model was more accurate and effective than the LSTM and Transformer models [9]. However, only the factor considered in this paper is time: if other factors are considered when training the model, the accuracy might be higher [9].

It is also common to see an integration of GRU with the attention model. Take Lee's work in 2022 as an example. Lee proposed a GRU-Attention deep neural network as a strategy reference for stock trading, and this study showed a significant improvement in prediction accuracy compared to other deep networks [10].

Despite using LSTM, GRU, Attention, and Transformer models, several other variations of LSTm models exist. For example, Qiu et al. proposed to predict stock price by using the WLSTM+Attention model [4]. The data is firstly processed by a wavelet transformer to make it more precise [4]. The prediction results were evaluated using S&P 500, DJIA, and HSI datasets. It was found that the WLSTM+Attention model outperformed several other models [4]. Another work by Kumar et al. proposed a method of forecasting the closing price of the stock market by using LSTM-TLBO [11]. Compared with the traditional LSTM model, TLBO focuses on execution speed, error frequency, and accuracy of results [11]. Research showed that TLBO outperforms other methods in optimizing stock price forecasts [11]. For large-scale processing of high-dimensional problems, TLBO is more effective in calculation [11]. Finally, Rajanand et al. proposed a DWCNN-SLSTM model, and they checked the performance on several baseline data sets by simply switching models while keeping all other network and training parameter constants [12]. As a result, they found that the proposed model is superior to the Transformer model in data sets in all performance indicators [12].

Although various deep learning models are available, we are still curious about which model can yield the most accurate predictions of different firms' stock prices. To achieve this goal, we will use the original LSTM, GRU, Attention Mechanism, and Transformer Models to predict the stock prices of different firms.

3 RELATED THEORIES

3.1 Long Short-Term Memory (LSTM)

The original LSTM was created in order to solve the problem of "long-term dependencies" and was proposed by Hochreiter and Schmidhuber in 1997 [6], which improves the memory capacity of standard circulating cells by bringing a "gate" into the cell. Then, the forget gate was introduced by Gers et al. in 2000 [13]. The following Figure 1 presents the inner connections of an LSTM with forget gates.

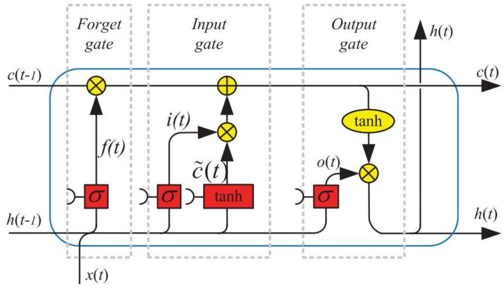


Figure 1 Inner Connections of an LSTM Cell [14]

Mathematically, we can present the inner structure of an LSTM unit with the following expressions:

$$f_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \tag{1}$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i) \tag{2}$$

$$\tilde{c}_t = \tanh\left(W_{\tilde{c}h}h_{t-1} + W_{\tilde{c}x}x_t + b_{\tilde{c}}\right) \tag{3}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{4}$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o) \tag{5}$$

$$h_t = o_t \bullet \tanh(c_t) \tag{6}$$

In Eq. (1)-(6), x_t, h_t, and c_t denote the input, the recurrent information, and the cell state, respectively. W_f, W_c, W_o, and W_i are the weights of forget gate, input gate, cell state, and output gate, and b is the bias. Further, f_t, i_t, and o_t are the activation functions used for output. The operator "·" is the pointwise multiplication of two vectors. When the value of a forget gate (f_t) is 1, it keeps the information. Alternatively, if the value of f_t is 0, it will delete the information.

3.2 Gated Recurrent Unit (GRU)

Although the LSTM cell is better than other standard recurrent cells, the additional parameters add computational burden. Hence, Cho et al. introduced the gated recurrent unit (GRU) in 2014 [15]. The Figure 2 below shows the details of the architecture and connections of a GRU cell:

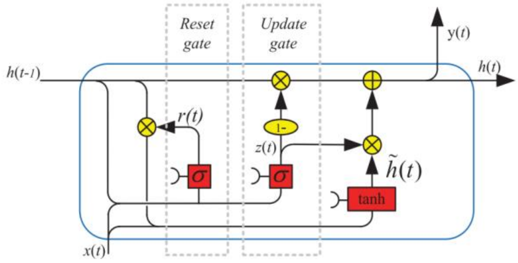


Figure 2 Architecture of a GRU Cell [14]

The following mathematical expressions are used to construct the GRU cells:

$$r_t = \sigma(W_{rh}h_{t-1} + W_{rx}x_t + b_r) \tag{7}$$

$$z_t = \sigma(W_{zh}h_{t-1} + W_{zx}x_t + b_z) \tag{8}$$

$$\tilde{h}_t = \tanh\left(W_{\tilde{h}h}(r_t \cdot h_{t-1}) + W_{\tilde{h}x}x_t + b_z\right) \tag{9}$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \tag{10}$$

In this model, Cho et al. (2014) integrate the forget gate and the input gate of the LSTM cell as an update gate to reduce the data needed to be computed [15]. However, since one gate is reduced in GRU, individual GRU cells are less potent than the original LSTM cells.

3.3 Attention Mechanism

Biological phenomena are essential in inspiring people to develop different powerful algorithms for deep learning models, and the attention mechanism is no exception. It is inspired by the study of human vision, which highlights the allocation of enough attention to information that is more important than others. Integrating the idea in stock price prediction, the attention mechanism is mainly used to predict stock prices by extracting news information. The attention mechanism was first implemented by Dzmitry as a soft research structure for French-English machine translation tasks [16]. With the expansion of the stock market and the demand for predicting stock prices, a recurrent neural network based on an attention mechanism is proposed to train financial news to predict stock prices [17]. The following Figure 3 represents the general structure of the attention mechanism:

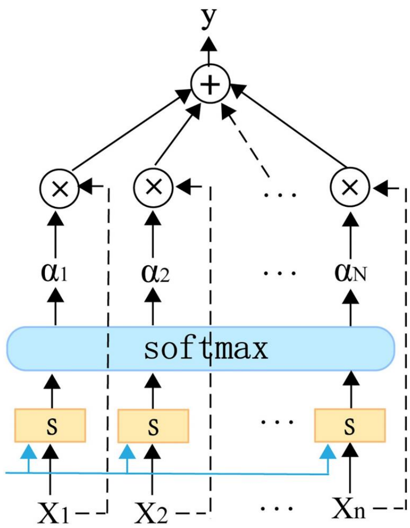


Figure 3 Architecture of the Attention Mechanism [4]

Dzmitry's study reveals the general mathematical expressions used to build an attention mechanism [16]:

$$e_{ij} = a(Q_i , K_j) \tag{11}$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_x} \exp(e_{ik})} \tag{12}$$

$$\text{Attention}(Q_i, K, V) = \sum_{j=1}^{T_x} \alpha_{ij} * V_j \tag{13}$$

In the equations above, Q_i is the query value corresponding to the i^{th} output element in the target. K denotes the key of all elements in the source, and, more specifically, K_j is the key of the j^{th} constituent element in the source. Moreover, V represents the value of all elements in the source, and similarly, V_j is the value of the j^{th} constituent element in the source. Lastly, T_x is the length of the source, and α is the calculation function of the correlation between Q and K_j.

3.4 Transformer Model

The Transformer model is a new generation of network architecture after Convolutional Neural Network (CNN) model proposed by Google [5]. It was initially used for natural language processing (NLP). However, due to its exact performance in downstream tasks, it is now widely used in computer vision to do tasks such as image classification, object detection, and image segmentation. The Transformer is developed based on the attention mechanism and thus is simpler and more violent than RNN. To be more specific, RNN obtains global information by recursion, whereas the transformer model based on the attention mechanism can obtain global information in only one step. The illustration below gives the architecture of a transformer model (Figure 4).

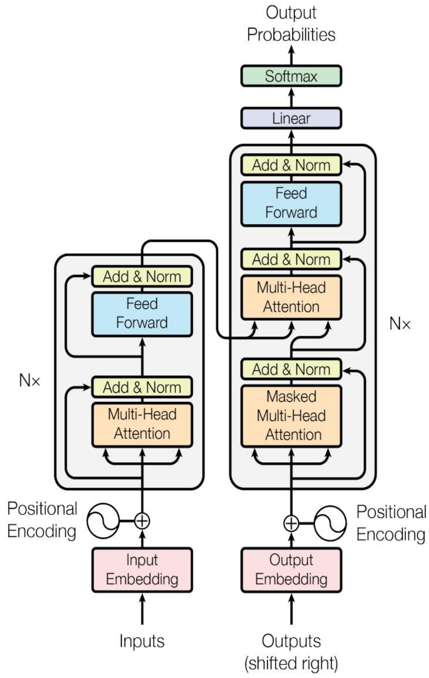


Figure 4 Architecture of the Transformer Model [5]

The following mathematical expressions give the essential idea of implementing a transformer model.

$$\mathrm{Attention}(Q,K,V)=\mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{14}$$

$$\mathrm{MultiHead}(Q,K,V)=\mathrm{Concat}(\mathrm{head}_1,\cdots,\mathrm{head}_h)W^O, \quad \text{where } \mathrm{head}_i=\mathrm{Attention}\left(QW_i^Q,KW_i^K,VW_i^V\right) \tag{15}$$

$$\mathrm{FFN}(x)=\max(0,xW_1+b_1)W_2+b_2, \tag{16}$$

$$\mathrm{PE}_{\mathrm{pos},2i}=\sin\left(\frac{\mathrm{pos}}{10000^{2i/d\mathrm{model}}}\right) \tag{17}$$

$$\mathrm{PE}_{\mathrm{pos},2i+1}=\cos\left(\frac{\mathrm{pos}}{10000^{2i/d\mathrm{model}}}\right) \tag{18}$$

In Eq. (14)-(18), Q is the query vector, K denotes the key vector, V is the value vector, and QK^T is a dot product operation that calculates the weight of attention for Q on V. The purpose of scaling the result by the square root of d_k is to avoid significantly large valuesin computation. Further, W^Q, W^K, and W^V are the three matrices computed during training. To introduce nonlinearity (ReLU activation function), FFN was added to increase the model's performance. Moreover, pos represents the position of the word, and *d*model is the dimension of the position vector, which equals to the dimension of the word encoding. Lastly, $i \in [0, d\mathrm{model})$ represents the i^{th} dimension of the position vector. The formula above gives us the *d*model vector at its corresponding pos position.

4 EXPERIMENTAL ANALYSIS

4.1 Data Collection

To compare different neural network models for stock price prediction, we collected stock prices of four firms from Yahoo Finance. Firms of selection include Boeing Co (NYSE: BA), General Electric Co (NYSE: GE), Coca-Cola Co (NYSE: KO), and Johnson & Johnson (NYSE: JNJ). The data collected includes the stock information of those enterprises on those trading days from Jan 2nd, 1962 to Nov 11th, 2017. The following table gives an example of the data collected:

Table 1 Sample Data For GE Stock Information Collected

Date	Close
1962-01-02	4.675709
1962-01-03	4.628796
1962-01-04	4.574063
1962-01-05	4.456780
1962-01-08	4.448961

In Table 1, Close refers to the closing price or the stock's final price on the corresponding trading day. We use the value of close to represent the price of all four stocks.

4.2 Data Processing

To better understand the stock price and for the sake of easy computation, we standardized the stock prices to a scale from –1 to 1. The following Figure 5 shows an example of normalized data of GE. Similar procedures are applied to the other three firms.

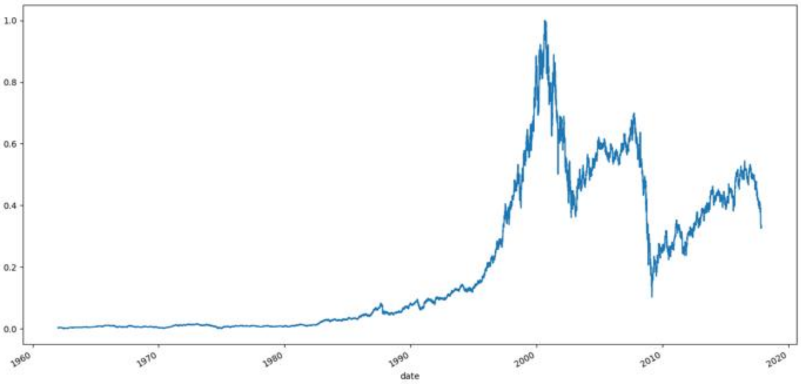


Figure 5 The General Trend of Normalzied GE Stock Price

Then, we move a lag window on the data set and classify the data into training sets, validation sets,s and testing sets. We got 9830 training data, 2810 validating data, and 1404 testing data in each data set.

4.3 Evaluation Metrics

Following the research done by Li et al., our study uses Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) to evaluate and compare the models [3]. The evaluation metrics can be calculated using the following formula:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y)^2 \tag{19}$$

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} \tag{20}$$

where N is the total number of samples, y_i hat is the value predicted by the model, and y_i is the expected value. The model is better and can yield more reliable outcomes with lower MSE and RMSE.

4.4 Discussion

In our study, we set the epoch of all the models to 100, and Table 2-5 gives the evaluation metrics of all the models with different stock prices.

Table 2 Evaluation Metrics for Predicting BA Prices

Model	MSE	RMSE
LSTM	0.0007	0.0266
GRU	0.0039	0.0625
Attention	0.0039	0.0625
Transformer	0.0020	0.0447

Table 3 Evaluation Metrics for Predicting GE Prices

Model	MSE	RMSE
LSTM	0.000051	0.0071
GRU	0.0009	0.0307
Attention	0.0002	0.0151
Transformer	0.0006	0.0251

Table 4 Evaluation Metrics for Predicting KO Prices

Model	MSE	RMSE
LSTM	0.0002	0. 0150
GRU	0.0003	0.0167
Attention	0.0007	0.0263

Transformer	0.0007	0.0263

Table 5 Evaluation Metrics for Predicting JNJ Prices

Model	MSE	RMSE
LSTM	0.0001	0.0139
GRU	0.0469	0.2165
Attention	0.0469	0.2165
Transforme r	0.0015	0.0390

All the trials agree that LSTM gives the most accurate predictions for stock prices because the MSE and RMSE of LSTM are the lowest among the four models in each trial. We are indecisive about which model is the second most accurate in predicting stock prices because the situation varies. As for predicting BA and JNJ prices, the transformer model is the second most accurate. However, when predicting GE prices, the attention mechanism becomes the second most accurate, whereas the GRU model gives the second most reliable predictions in the case of KO price prediction.

To further our understanding of price prediction using LSTM, we plot the predicted stock price of LSTM with the actual value on the same graph, respectively, with different firms. Figure 6-9 shows the outcomes of LSTM according to the four firms of selection:

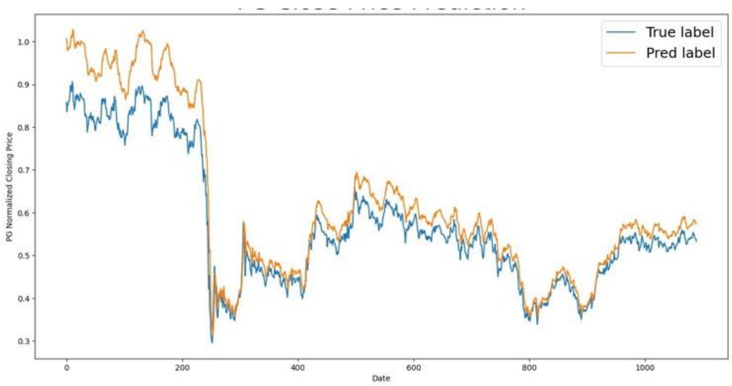


Figure 6 LSTM Prediction of BA Stock Price

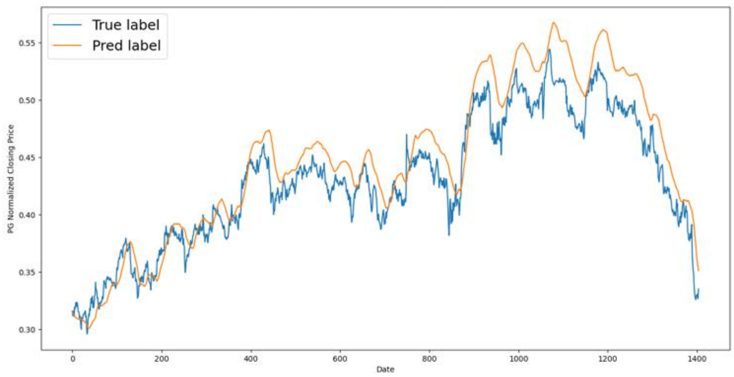


Figure 7 LSTM Prediction of GE Stock Price

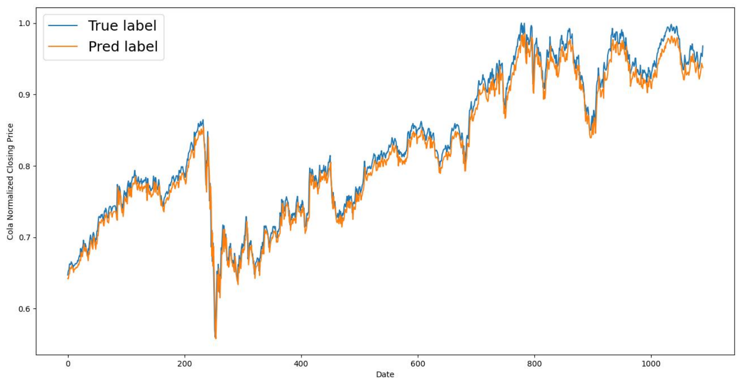


Figure 8 LSTM Prediction of KO Stock Price

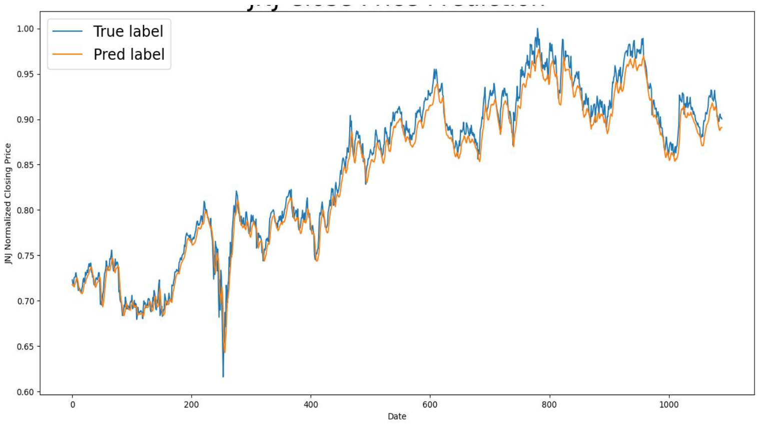


Figure 9 LSTM Prediction of JNJ Stock Price

All the plots suggest that LSTM is suitable for predicting the general trends of the stock price and can signal sudden but significant increases or decreases. In other words, when the stock price increases, the model returns an increasing trend as predicted. Moreover, when the actual data tends to decrease, the model also agrees with the tendency of the real data.

However, when predicting the BA price, the model shows relatively poor fitness at the beginning of the prediction, but as time goes on, the prediction becomes better and better. As for the GE price prediction, the model can only predict the general trend of the price but needs more precision.

To improve the consistency of accuracy and precision of prediction, we could consider adding more training data for the model to predict the stock price with higher accuracy and precision even at the very beginning of prediction. Moreover, we can integrate other factors into our model, such as stock news or fluctuations of other stock prices, so that the model can better understand the stock market and make better predictions based on that additional information. Lastly, we could integrate different models, and thus not only can the model yield more accurate results, but also we could save time and the amount of original data needed.

5 CONCLUSION

This paper explored the history of using deep learning models to predict the stock market. Then we compared the accuracy of stock predictions of four mainstream deep learning models: LSTM, GRU, Attention, and Transformer. We used MSE and RMSE as our evaluation metrics and found that LSTM provides the most accurate prediction. However, we also found that the model lacks consistency in predicting the stock prices of different firms. To improve, we suggest adding more training data, introducing additional factors, and integrating different models so that the model can understand the stock market better and yield more accurate and precise predictions.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

ACKNOWLEDGMENT

L. L. W. Y. sincerely thanks the anonymous author for his help and dedication.

REFERENCES

[1] Q Ding, S Wu, H Sun, et al. Hierarchical Multi-Scale Gaussian Transformer for Stock Movement Prediction. Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, Yokohama, Japan: International Joint Conferences on Artificial Intelligence Organization, 2020: 4640–4646. DOI: 10.24963/ijcai.2020/640.

[2] D Nelson, A Pereira, R de Oliveira. Stock market's price movement prediction with LSTM neural networks. 2017: 1419–1426. DOI: 10.1109/IJCNN.2017.7966019.

[3] H Li, Y Shen, Y Zhu. Stock price prediction using attention-based multi-input LSTM. Journal of Financial Data Science. 2022, 4(3): 45-60. DOI: 10.1016/j.jfds.2022.05.002.

[4] J Qiu, B Wang, C Zhou. Forecasting stock prices with long-short term memory neural network based on attention mechanism. PLoS One, 2020, 15(1): e0227222. DOI: 10.1371/journal.pone.0227222.

[5] A Vaswani, N Shazeer, N Parmar, et al. Attention Is All You Need. 2017. http://arxiv.org/abs/1706.03762.

[6] S Hochreiter, J Schmidhuber. Long Short-Term Memory. Neural Computation, 1997, 9(8): 1735–1780.

[7] U Gupta, V Bhattacharjee, P Bishnu. StockNet - GRU based Stock Index Prediction. Expert Systems with Applications, 2022, 207: 117986. DOI: 10.1016/j.eswa.2022.117986.

[8] C Wang, Y Chen, S Zhang, et al. Stock market index prediction using deep Transformer model. Expert Systems with Applications, 2022, 208: 118128. DOI: 10.1016/j.eswa.2022.118128.

[9] S Zhang, H Zhang. Prediction of Stock Closing Prices Based on Attention Mechanism. 2020 16th Dahe Fortune China Forum and Chinese High-educational Management Annual Academic Conference (DFHMC), Zhengzhou, China: IEEE, 2020. DOI: 10.1109/DFHMC52214.2020.00053.

[10] M C Lee. Research on the Feasibility of Applying GRU and Attention Mechanism Combined with Technical Indicators in Stock Trading Strategies. Applied Sciences, 2022, 12(3). DOI: 10.3390/app12031007.

[11] K Kumar M, V R C, I C Kumari P, et al. Stock Price Prediction using LSTM and TLBO. 2023 IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), 2023: 1–5. DOI: 10.1109/ICICACS57338.2023.10100074.

[12] A Rajanand, P Singh. Stock Price Prediction using Depthwise Pointwise CNN with Sequential LSTM. 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2023: 82–86. DOI: 10.1109/ICAAIC56838.2023.10140728.

[13] F A Gers, J Schmidhuber, F Cummins. Learning to forget: continual prediction with LSTM. Neural Comput, 2000, 12(10): 2451–2471. DOI: 10.1162/089976600300015015.

[14] Y Yu, X Si, C Hu, J Zhang. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures," Neural Computation, 2019, 31(7): 1235–1270. DOI: 10.1162/neco_a_01199.

[15] K Cho, B van Merrienboer, C Gulcehre, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. 2014. DOI: 10.48550/arXiv.1406.1078.

[16] D Bahdanau, K Cho, Y Bengio. Neural Machine Translation by Jointly Learning to Align and Translate. arXiv, 2016. http://arxiv.org/abs/1409.0473.

[17] H Liu. Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network. 2018. DOI: 10.48550/arXiv.1811.06173.

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951
Online ISSN: 3007-696X
DOI: https://doi.org/10.61784/jtfe3064

HIGH-QUALITY DEVELOPMENT FOR RURAL CULTURAL INDUSTRIES IN THE GUANGDONG-HONG KONG-MACAO GREATER BAY AREA FROM THE PERSPECTIVE OF NEW QUALITY PRODUCTIVITY FORCES

TingTing Zheng, ZiChen Wang*
School of Management, Guangzhou Huashang College, Guangzhou 511300, Guangdong, China.
Corresponding Author: ZiChen Wang, Email: wzcwqljk@163.com

Abstract: This study examines the high-quality development of rural cultural industries in the Guangdong-Hong Kong-Macao Greater Bay Area from the perspective of new-type productive forces. Findings reveal a gradient distribution of cultural resources across the region, challenges to industrial advancement include limited application of new-type productive forces, shortages of talent and capital, inadequate infrastructure, weak technological capabilities, and insufficient cross-border policy coordination. To address these challenges, this paper constructs an analytical framework centered on "new factors-integration models-policy coordination" and empirically proposes three development pathways. These pathways aim to provide theoretical support and practical solutions for achieving urban-rural integration and rural revitalization: First, digital empowerment, involving infrastructure development and the creation of digital cultural and creative products; Second, creative activation, focusing on cultivating design talent and building intellectual property systems; Third, cross-border coordination, aimed at breaking down policy barriers and facilitating project implementation.
Keywords: New quality productivity forces; High-quality development; Guangdong-Hong Kong-Macao Greater Bay Area

1 INTRODUCTION

The 2025 Central Government Document No. 1 explicitly proposes "empowering rural industrial revitalization with new-quality productive forces," integrating new elements such as digital technology and creative design into the rural economic system, thereby providing strategic guidance for the transformation of rural cultural industries. As one of China's most open and economically dynamic regions, the Guangdong-Hong Kong-Macao Greater Bay Area possesses triple advantages: Hong Kong region and Macao region's international cultural resources, the Pearl River Delta's manufacturing foundation, and Western Guangdong's ecological and cultural heritage. Resources such as Hong Kong region's cultural and creative design, Shenzhen's digital technology, Jiangmen's watchtower intangible cultural heritage, and Zhaoqing's Duan inkstone culture urgently require synergistic activation through new quality productive forces. However, the region's rural cultural industries currently face three core challenges: First, superficial industrial integration, with most projects confined to "culture + tourism" sightseeing models lacking deep integration with agriculture and services. Second, weak market competitiveness, characterized by severe homogenization of cultural products in peripheral areas and a shortage of digitally empowered premium brands. Third, fragmented policy support, marked by insufficient coordination among the three regions, with Hong Kong region and Macau capital facing institutional barriers like land use and taxation when entering mainland rural cultural industries.

This study introduces the theory of "new-quality productive forces" into rural cultural industry research, expanding its application boundaries in the cultural economy. It constructs an analytical framework of "new factors—integration models—policy coordination" to enrich the theoretical system for high-quality development of rural cultural industries. Based on empirical data from the Greater Bay Area, it distills a development model of "digital empowerment + creative activation + cross-border coordination." This provides scientific basis for governments to formulate differentiated policies, assists peripheral rural cultural industries in enhancing added value, and promotes common prosperity.

2 THEORETICAL FOUNDATIONS AND LITERATURE REVIEW

The digital transformation of cultural industries represents a defining feature of their leapfrog development in the new era, embodying the fundamental role of new-quality productive forces in empowering high-quality growth. "New-quality productive forces" denote a form of productivity driven by technological innovation, supported by new elements such as digital technology, artificial intelligence, and creative design, and characterized by high technological content and value-added potential—transcending traditional growth pathways [1]. Its core characteristics include: first, factor innovation, where digital technology and creative talent replace traditional factors like land and labor; second, integration and penetration, driving cross-industry convergence through technological empowerment; and third, value multiplication, enhancing product value-added through creative design.

In the digital economy era, new-quality productive forces inject fresh momentum into the high-quality development of rural cultural industries through the enabling mechanisms of technological innovation, institutional reform, and industrial transformation [2]. High-quality development of rural cultural industries must fulfill "three-dimensional value": First, economic value, achieving industrial appreciation and increased farmer income—such as the Jiangmen Diaolou Digital Cultural and Creative Project, which boosted villagers' average annual income by 25%; Second, cultural value, which involves preserving and revitalizing intangible cultural heritage resources. Digital technology creates authentic rural cultural scenes, showcasing vibrant and diverse rural cultural characteristics while establishing a comprehensive digital rural resource repository[3]. Third, social value, which facilitates the flow of urban-rural resources. For instance, Shenzhen Dapeng New District's "Intangible Cultural Heritage Digital Museum" attracted 120 urban designers to engage with rural communities.

New-type productive forces, crystallizing the continuous enhancement of productive elements' quality, represent a more advanced form of productivity. Its core driving force stems from innovation, with education as its foundational support. Talent serves as the bridge connecting all elements, while industries provide a broad implementation platform[4]. By cultivating high-caliber rural revitalization talent through education, driving industrial convergence via technological innovation, strengthening social service support, and establishing collaborative platforms, new vitality is injected into rural revitalization [5]. This study extends these principles to rural cultural industries: digital technology addresses "resource activation," creative design tackles "product value enhancement," and cross-border capital resolves "funding shortages."

Existing research primarily focuses on: First, regional disparities. Development levels vary significantly among cities in the Guangdong-Hong Kong-Macao Greater Bay Area, resulting in a "core-periphery" gradient distribution for rural cultural industries within the region. Core areas predominantly feature digital cultural and creative industries, while peripheral areas focus on intangible cultural heritage tourism [6]. Second, industrial integration. In recent years, Guangdong's industrial and regional policies have undergone substantial adjustments, increasingly planning industrial integration between relocating enterprises and receiving areas from perspectives of coordinated planning, complementary advantages, and shared development [7]. Third, policy coordination. The Greater Bay Area emphasizes enterprises as key cultural development actors primarily to diversify cultural leadership. This approach enables enterprises to play a more prominent role in fulfilling their social responsibilities while allowing the market to participate collaboratively in the region's cultural development, thereby leveraging market mechanisms in cultural advancement [8].

3 CURRENT STATUS OF RURAL CULTURAL INDUSTRY DEVELOPMENT

3.1 Gradient Distribution of Resource Endowments

Rural cultural resources in the Guangdong-Hong Kong-Macao Greater Bay Area exhibit a typical pattern of "creative dominance in core areas and resource concentration in peripheral regions," which can be categorized into three tiers:

The core region primarily includes Hong Kong region, Macao region, Guangzhou, and Shenzhen. These cities leverage digital creativity and modern design as their core development resources. For instance, Hong Kong region continuously provides policy support through its "Create Hong Kong" initiative, which invests HK$200 million annually to support cultural and creative enterprises. Shenzhen, meanwhile, hosts a vast cluster of cultural and creative enterprises, with incomplete statistics indicating the number has reached 150,000.

The secondary zone primarily encompasses Foshan, Dongguan, and Zhongshan. Leveraging their robust manufacturing foundations, these cities actively develop the "culture + manufacturing" integration model, forming distinctive industrial characteristics. Take Xiqiao Town in Foshan as an example: its iconic Xiangyunshan silk industry has achieved an annual output value of 1 billion RMB.

Peripheral regions include Jiangmen, Zhaoqing, and Huizhou. These areas possess exceptionally rich intangible cultural heritage and ecological resources. For instance, Jiangmen's Kaiping Diaolou (watchtower houses), a UNESCO World Heritage site, currently number 1,833 structures. Huizhou's Longmen Folk Paintings, a national intangible cultural heritage, see an average annual production of approximately 2,000 pieces locally. However, the current utilization rate of these abundant resources remains insufficient, generally below 20%.

3.2 Stages of Industry Development

Based on a comprehensive assessment of industrial integration and value-added potential, the development of rural cultural industries in the Guangdong-Hong Kong-Macao Greater Bay Area can be divided into three progressive stages:

At the primary stage, this model is concentrated in peripheral areas, centered on the traditional "culture + tourism" sightseeing approach, with cultural elements appearing rather singularly within tourism products. For instance, cultural product revenue accounts for only 10% of total income at Zhaoqing's Dinghu Mountain scenic area, indicating that the penetration and value conversion of cultural elements within the region's tourism industry chain remain underdeveloped.

Progressing to the intermediate stage, industrial integration deepens, forming preliminary "culture + manufacturing" models in secondary zones. Dongguan's Chashan Town exemplifies this by combining the cultural IP of the "Nanshe Ming-Qing Ancient Village" with creative product development, achieving initial conversion of cultural value into economic value. Related industries now generate an annual output value of 200 million yuan.

At the advanced stage of industrial integration, core areas leverage technological, capital, and talent advantages to pioneer a deep integration model of "culture + digital + finance." Shenzhen's Dapeng New District exemplifies this with its "Intangible Cultural Heritage Digital Museum." By revitalizing intangible cultural heritage resources through digital technology, the project attracts over 2 million annual visitors, significantly boosting income growth by 35% in surrounding industries like homestays.

3.3 Current Application of New Quality Productivity in Rural Cultural Industries of the Greater Bay Area

Based on interviews with industry professionals in the Greater Bay Area, the application of new quality productive forces in rural cultural industries exhibits a pattern of "deep penetration in core areas and shallow application in peripheral areas":

Core areas (Hong Kong region, Macao region, Guangzhou, Shenzhen) exhibit a digital technology penetration rate as high as 65%, primarily concentrated in digital exhibitions and e-commerce sales. For instance, Hong Kong region's Yuen Long District revitalizes rural cultural resources through immersive technologies like "VR Farming Experiences," while Shenzhen's Dafen Oil Painting Village leverages live-streaming e-commerce to achieve 40% online sales, significantly enhancing cultural product reach and transaction efficiency.In contrast, peripheral areas exhibit only about 25% digital technology penetration, with relatively basic applications primarily focused on digitizing foundational services. For instance, the Kaiping Diaolou in Jiangmen improved visitor experiences through an online ticketing system, yet overall digital depth and application breadth remain significantly underdeveloped.

Core regions achieve 70% creative design participation, forming a complete chain of "cultural IP cultivation—product innovation—value enhancement." Taking Shenzhen as an example, its annual "Bao'an Cultural and Creative IP Design Competition" attracts numerous design talents, successfully producing 100 original cultural IPs and driving the transformation of traditional symbols into modern cultural and creative products. In contrast, most cultural products in peripheral regions still rely on traditional craftsmanship, with insufficient application of innovative design.For instance, Zhaoqing Duan inkstones, a national intangible cultural heritage, typically undergo product design updates only every five years or more, struggling to adapt to market shifts. Thus, the driving force of creative design for industrial upgrading remains underutilized.

Hong Kong region and Macao region capital participation in rural cultural industries within the Greater Bay Area remains low, accounting for less than 10% of total projects and primarily concentrated in core areas. For instance, Hong Kong region cultural enterprises invested in Shenzhen's Dapeng New District "Marine Cultural Creative Park," leveraging cross-border collaboration to integrate resources.Projects involving Hong Kong region and Macao region capital in peripheral areas account for less than 3% of the total. The core obstacle lies in land policy restrictions—for example, collective construction land in some villages in Jiangmen cannot be directly transferred to Hong Kong region and Macao region enterprises. This results in institutional constraints on cross-border investment during the land acquisition process, hindering the cross-regional flow and integration of resources.

3.4 Core Bottlenecks in New Quality Productivity Application

3.4.1 Factor bottlenecks: shortage of digital talent and creative capital

Talent shortages manifest as severe underprovision of digital technology personnel for rural cultural industries in peripheral areas. For instance, digital technology professionals constitute less than 5% of rural cultural industry workers in these regions. Taking Huizhou Longmen's farmer paintings as an example, only three individuals among numerous local intangible cultural heritage inheritors possess design software operation skills, failing to meet digital creation and operational demands.

Capital constraints are evident in the significant financing gap between peripheral and core regions for cultural projects. Survey data indicates that the average funding for rural cultural projects in peripheral areas is merely 500,000 yuan, whereas comparable projects in core regions secure an average of 5 million yuan. This insufficient capital supply directly hampers the implementation and application of new-quality productive forces in peripheral regions.

3.4.2 Technical bottlenecks: weak digital infrastructure and application capabilities

Lagging infrastructure deprives peripheral areas of hardware support for digital technology adoption. Currently, 5G network coverage in rural peripheral areas stands at only 40%, far below the 90% coverage in core regions. Insufficient network bandwidth and stability fail to meet the technical demands of new productive forces applications such as digital exhibitions and remote collaboration.

Insufficient application capabilities manifest at the enterprise operational level. Most cultural enterprises in peripheral areas lack professional digital operations teams. For instance, the online marketing efforts for the Kaiping Diaolou Scenic Area in Jiangmen are managed by only one part-time staff member. This results in limited and single-dimensional digital promotion methods, hindering the full realization of digital technology's potential to enhance industrial efficiency.

3.4.3 Institutional bottlenecks: insufficient cross-border policy coordination

Policy discrepancies increase the institutional costs for Hong Kong region and Macao region capital entering mainland rural cultural industries. When undertaking rural cultural projects in the mainland, Hong Kong region and Macao region investors must undergo additional procedures like "foreign investment filing," with approval cycles lasting up to three months. The complexity of these policy processes slows project implementation.

Lack of unified standards hinders the efficient flow of cultural resources across the three regions. Differences in cultural product quality standards and intellectual property protection rules—such as Hong Kong region cultural and creative products requiring re-inspection upon entering the mainland market—increase transaction costs for cross-regional cooperation, constraining the cross-border integration and optimal allocation of new productive forces.

4 PATHWAYS FOR HIGH-QUALITY DEVELOPMENT

4.1 Digital Empowerment Pathway: Building a Digital Ecosystem for Rural Cultural Industries

Leveraging digital technology as the core driver, promote the digital transformation of rural cultural industries by enhancing infrastructure, cultivating operational capabilities, and developing cultural and creative products.

Enhance digital infrastructure: Adopt a "core-area radiation + peripheral-area gap-filling" model to extend 5G network coverage from core areas like Shenzhen and Guangzhou to peripheral regions such as Huizhou and Jiangmen, aiming to increase 5G coverage in rural peripheral areas to 80% by 2025–2027.Simultaneously deploy AR guide devices and digital display screens at key cultural sites like watchtower villages and folk painting villages. For instance, the Jiangmen Watchtower Scenic Area plans to install 100 pairs of AR guide glasses to enhance visitor interaction experiences.

Cultivating Digital Operational Capabilities: Through coordinated efforts of "government training + corporate collaboration + talent deployment," annually organize the "Greater Bay Area Rural Cultural Industry Digital Operations Training Program" to cultivate 100 professionals. Facilitate partnerships between core-area digital enterprises like Tencent and the Jiangmen Watchtower Scenic Area to co-develop WeChat mini-programs. Implement the "Digital Talent Rural Deployment Plan," encouraging technical personnel from core areas to serve in peripheral-area enterprises for 1–2 years while retaining original employer benefits plus special subsidies.

Developing Digital Cultural Products: Guided by "traditional culture + digital technology + modern aesthetics," transform intangible cultural heritage resources like Longmen Peasant Paintings and Duan inkstones into digital collectibles, targeting ¥5 million in sales for peasant painting digital collectibles by 2024. Establish the Zhaoqing Duan Inkstone Metaverse Experience Hall, offering immersive tours of traditional craftsmanship.Design hardware products like Jiangmen Diaolou-themed smart speakers, with projected sales of 100,000 units by 2025, integrating cultural IP with consumer electronics.

4.2 Creative Activation Pathway: Building a Creative Ecosystem for Rural Cultural Industries

Using creative design as the core link, enhance the innovative vitality of rural cultural industries by cultivating local talent, introducing external resources, and building IP systems.

Cultivating Local Creative Talent: Establishing a three-dimensional system integrating "school education + social training + master studios," Jiangmen Polytechnic has launched a "Watchtower Cultural and Creative Design" program. The Shenzhen Graphic Design Association trains 50 designers annually from marginal areas. Hong Kong region designer Alan Chan collaborates with Zhaoqing Duan inkstone artisans to establish a studio, integrating traditional craftsmanship with modern design. This initiative has boosted artisans' monthly income from 3,000 to 12,000 RMB.

Introducing External Creative Resources: Hosting the "Greater Bay Area Rural Cultural Creative Design Competition" with a 1 million RMB prize pool to attract proposals from core areas and Hong Kong region/Macao region; signing 10 creative projects totaling 500 million RMB in investment at a Shenzhen investment conference; providing free design consultations for peripheral enterprises through Macao region's cultural and creative platforms, serving 20 projects in 2024 to elevate product design standards.

Establishing a Creative IP System: Systematically cataloging cultural symbols like watchtowers and peasant paintings to develop IPs such as "Watchtower Guardians" and "Peasant Painting Babies"; establishing the "Peripheral Rural Cultural IP Operations Center" to manage IP licensing and commercial development; creating stationery, toys, and other derivatives around core IPs, with projected 2025 sales reaching 100 million yuan to build sustainable IP monetization capacity.

4.3 Cross-Border Collaboration Pathway: Building a Cross-Border Collaborative Ecosystem for Rural Cultural Industries

Leveraging Hong Kong region and Macao region resources as key pillars, promote deep integration of rural cultural industries across the three regions through policy coordination, factor mobility, and project collaboration.

Establish a cross-border policy coordination mechanism: Develop the "Greater Bay Area Rural Cultural Industry Cross-Border Cooperation Negative List" to clarify prohibited and restricted projects; create a "cross-border approval green channel" to reduce the approval cycle for Hong Kong region and Macao region capital projects from 3 months to 7 working days; establish a tri-regional product standards committee to unify quality specifications and intellectual property protection rules, eliminating trade barriers.

Develop a cross-border factor mobility platform: Launch the "Greater Bay Area Rural Cultural Industry Cross-Border Creative Platform" to integrate creative talent, capital, and market resources across the three regions; establish a factor exchange center offering cross-border talent, technology, and copyright transaction services; provide free exhibition booths for enterprises from peripheral areas at the Hong Kong International Cultural and Creative Expo, with 15 enterprises securing over US$20 million in overseas orders by 2024.

Launch cross-border cooperation projects: Construct the "Hong Kong-Macao Cultural and Creative Industrial Park" in Kaiping, Jiangmen, targeting 50 enterprises by 2025; develop the "Hong Kong Yuen Long–Shenzhen Dapeng–Jiangmen Diaolou" cross-border tourism route to connect cultural resources across the three regions; jointly create "Greater Bay Area Rural Cultural IPs," with plans to launch 10 collaborative IPs by 2025 and promote them internationally through cross-border channels.

4.4 Policy Support Pathway: Establishing a Policy Framework for Rural Cultural Industries

Leverage fiscal, land, and talent policies as pillars to provide full-cycle support for the development of new quality productive forces.

Fiscal Support Policies: Establish the "Greater Bay Area Rural Cultural Industry New Quality Productivity Special Fund," investing 500 million yuan annually to support infrastructure and creative projects in peripheral areas; implement a "three-year exemption, two-year reduction" tax incentive for enterprises in peripheral areas (exempting corporate income tax for the first three years, then halving it for the next two); encourage financial institutions to develop "Digital Cultural and Creative Loans" with a maximum limit of 10 million yuan, prioritizing support for technology application and IP development.

Land Support Policy: Pilot the "spot land supply" model, such as the Huizhou Longmen Farmers' Painting Project, which allocated only 5 mu of land while preserving 95 mu of ecological green space; permit direct transfer of collective construction land in peripheral areas to Hong Kong region and Macao region enterprises, such as the Kaiping pilot in Jiangmen where Hong Kong region and Macao region enterprises acquired usage rights through land transfers; establish an urban-rural construction land quota trading platform to ensure land supply for key projects.

Talent Support Policy: Provide monthly living subsidies of 5,000 yuan for digital and creative talents deployed to peripheral areas, with a maximum duration of three years; construct "Rural Cultural Talent Apartments" offering move-in-ready accommodations; enable children of core-area talents to attend key schools in peripheral areas—e.g., Shenzhen-deployed talents' children may enroll in Jiangmen's top institutions—eliminating concerns about talent mobility.

5 CONCLUSION

This study employs a mixed-methods approach to reveal the practical challenges and breakthrough pathways for high-quality development of rural cultural industries in the Guangdong-Hong Kong-Macao Greater Bay Area.

Research indicates that new-quality productive forces are key to resolving current industry problems—digital technology addresses resource activation, creative design enhances product value-added, and cross-border capital compensates for funding shortages. Establishing a three-dimensional pathway system of “digital empowerment + creative activation + cross-border collaboration,” supported by complementary policies, can effectively propel the transformation of the Bay Area's rural cultural industries from “resource dependency” to “innovation-driven” models, achieving integrated urban-rural development and shared prosperity. Future efforts should further strengthen policy coordination among the three regions and improve mechanisms for factor mobility, providing more robust institutional safeguards for empowering rural revitalization through new quality productive forces.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

Funding Project: Guangzhou Huashang College 2024 Young Scholars Research Project “Research on Pathways for High-Quality Development of Rural Cultural Industries in the Guangdong-Hong Kong-Macao Greater Bay Area under the Context of Rural Revitalization” (2024HSQX013)

REFERENCES

[1] Xiang Yong. Value Construction of New Quality Productivity and High-Quality Development of Digital Cultural Industries. Jiangsu Social Sciences, 2024(05): 68-76+242.

[2] Rao Hongmei. Empowering High-Quality Development of Rural Cultural Industries Through New Quality Productivity: Driving Mechanisms and Implementation Pathways. Journal of Liaoning Agricultural Vocational and Technical College, 2025, 27(05): 29-33.

[3] Yuan Yuan, Fan Jianhua. The Internal Logic and Pathways of Digitalization in Promoting High-Quality Development of Rural Cultural Industries. Theory Monthly, 2023(08): 83-92.

[4] Li Xuelin, Liu Meiyi. New Quality Productivity Boosting the Revitalization of Rural Cultural Industries: Functional Positioning and Path Selection. Journal of Shaanxi Administrative College, 2024, 38(03): 109-114.

[5] Mao Xinyuan, Ren Zhaochang, Chen Weihao, et al. Theoretical Logic and Implementation Pathways for Empowering Modernization of Agriculture and Rural Areas in Yunnan Province through New Quality Productivity. Rural Economy and Technology, 2024, 35(12): 125-129.
[6] Zhang Alan, Chen Yougen, Lin Weiqiang. Innovation in Rural Revitalization Models in the Guangdong-Hong Kong-Macao Greater Bay Area from the Perspective of Urban-Rural Cultural Tourism Integration. Rural Economy and Technology, 2023, 34(15): 125-128.
[7] Cao Zongping. On the Strategic Integration of Rural Revitalization and the Development of the Guangdong-Hong Kong-Macao Greater Bay Area. Open Guide, 2023(03): 48-58.
[8] Qiu Daidong. Policy Preferences in Cultural Development of the Guangdong-Hong Kong-Macao Greater Bay Area: A Discussion Centered on Guangdong Province. Comparative Studies on Cultural Innovation, 2023, 7(04): 130-134.

Journal of Trends in Financial and Economics
Print ISSN: 3007-6951
Online ISSN: 3007-696X
DOI: https://doi.org/10.61784/jtfe3065

PREDICTIVE ANALYTICS FOR TRANSFER PRICING AND ITS REGULATORY IMPLICATIONS

JiaYing Chen[1*], Pan Li[2], Huijie Fan[3]
[1]Cornell University, Ithaca 14850, New York, USA.
[2]University of Hull, Hull HU6 7RX, East Riding of Yorkshire, UK.
[3]Hunan Agricultural University, Changsha 410128, Hunan, China.
Corresponding Author: JiaYing Chen, Email: jc2744@cornell.edu

Abstract: Transfer pricing (TP) has become increasingly complex in the era of globalization, requiring multinational enterprises (MNEs) to establish arm's length prices for intercompany transactions across jurisdictions. Traditional transfer pricing methodologies, while established through decades of regulatory practice, face significant challenges in addressing the complexity and volume of modern cross-border transactions. The emergence of predictive analytics (PA) and machine learning (ML) techniques offers transformative potential for enhancing transfer pricing determination, documentation, and compliance. This review examines the application of predictive analytics in transfer pricing contexts, exploring how artificial intelligence (AI), big data analytics (BDA), and advanced statistical methods are reshaping both corporate tax planning strategies and regulatory enforcement mechanisms. The regulatory implications of these technological advances are profound, raising questions about data transparency, algorithmic accountability, and the evolution of arm's length principle (ALP) interpretation. This paper synthesizes current research on predictive modeling approaches including neural networks (NN), random forests (RF), gradient boosting machines (GBM), and support vector machines (SVM) applied to comparable company selection, profit allocation, and risk assessment. We examine how tax authorities worldwide are deploying similar technologies for audit selection and compliance monitoring, creating both opportunities and challenges for MNEs navigating increasingly data-driven regulatory environments. The review addresses critical implementation considerations including data quality requirements, model interpretability standards, and the alignment of predictive systems with existing legal frameworks under Organisation for Economic Co-operation and Development (OECD) guidelines and local regulations. Findings indicate that while predictive analytics significantly improves accuracy and efficiency in transfer pricing processes, successful implementation requires careful attention to regulatory acceptability, documentation standards, and cross-functional integration between tax, finance, and data science teams.
Keywords: Transfer pricing; Predictive analytics; Machine learning; Tax compliance; Regulatory implications; Arm's length principle; Multinational enterprises; Artificial intelligence; OECD guidelines; BEPS

1 INTRODUCTION

Transfer pricing (TP) represents one of the most challenging areas of international taxation, governing the pricing of transactions between related entities within multinational enterprises (MNEs) across different tax jurisdictions. The fundamental principle underlying transfer pricing regulation is the arm's length principle (ALP), which requires that intercompany transactions be priced as if they occurred between independent parties under comparable circumstances [1]. This principle, codified in Article 9 of the Organisation for Economic Co-operation and Development (OECD) Model Tax Convention and embedded in domestic legislation across more than 60 countries, aims to prevent profit shifting and ensure appropriate tax revenue allocation among jurisdictions [2]. However, the practical application of ALP has grown increasingly complex due to the expansion of global value chains, the proliferation of intangible assets, and the digitalization of business models that challenge traditional TP methodologies [3].

The compliance burden associated with TP has escalated dramatically in recent years, driven by enhanced regulatory scrutiny following the OECD Base Erosion and Profit Shifting (BEPS) initiative and the introduction of country-by-country reporting (CbCR) requirements [4]. MNEs now face extensive documentation obligations, requiring detailed functional analysis, economic analysis, and benchmarking studies to support their intercompany pricing policies [5]. Traditional approaches to TP analysis rely heavily on manual processes, expert judgment, and retrospective application of established methods such as the comparable uncontrolled price (CUP) method, resale price method (RPM), cost plus method (CPM), transactional net margin method (TNMM), and profit split method (PSM) [6]. These conventional methodologies, while theoretically sound, suffer from significant limitations including subjectivity in comparable selection, limited data availability, difficulty in adjusting for differences between controlled and uncontrolled transactions, and challenges in addressing unique value creation aspects of modern business models [7].

The emergence of predictive analytics (PA) and machine learning (ML) technologies offers transformative potential for addressing these limitations and enhancing TP practices. PA encompasses a range of statistical and computational techniques designed to identify patterns in historical data and generate predictions about future outcomes or unknown parameters [8]. When applied to TP contexts, PA can improve the accuracy of comparable company identification, enhance the precision of arm's length range determination, enable real-time monitoring of TP outcomes, and provide more robust support for documentation and defense positions [9]. Advanced ML algorithms including neural networks (NN), random forests (RF), gradient boosting machines (GBM), and support vector machines (SVM) have demonstrated superior performance compared to traditional statistical methods in handling high-dimensional data, capturing non-linear relationships, and automating complex pattern recognition tasks [10].

The adoption of PA in TP practice extends beyond mere technical implementation and carries significant regulatory implications that warrant careful examination. Tax authorities worldwide are simultaneously deploying similar technologies for audit selection, risk assessment, and compliance monitoring, fundamentally altering the dynamics of tax administration and enforcement [11]. The use of algorithmic decision-making in both corporate tax planning and government oversight raises critical questions about transparency, interpretability, and the appropriate evolution of established legal principles in light of technological capabilities [12]. Furthermore, the integration of big data analytics (BDA) with TP processes creates new considerations regarding data privacy, cross-border data flows, and the evidentiary standards applicable to algorithmically-generated analyses in tax dispute resolution contexts [13].

This review paper examines the current state of research and practice regarding the application of PA to TP determination and compliance, with particular emphasis on the regulatory implications of these technological developments. The analysis addresses several key research questions that have emerged as central to understanding the transformative potential and limitations of PA in this domain. First, what specific PA methodologies have proven most effective for different aspects of TP analysis, and what are their respective strengths and limitations? Second, how are tax authorities incorporating PA into their compliance and enforcement strategies, and what implications does this have for MNEs' approach to TP risk management? Third, what regulatory and legal frameworks are emerging to govern the use of algorithmic analyses in TP contexts, and how do these frameworks balance innovation with established principles of tax law? Fourth, what implementation challenges do organizations face when deploying PA for TP purposes, and what best practices have emerged for addressing these challenges? The structure of this paper proceeds through comprehensive literature review, examination of specific PA methodologies, analysis of regulatory implications, and discussion of implementation considerations.

2 LITERATURE REVIEW

The intersection of PA and TP represents an emerging research domain that has gained substantial attention since 2019, reflecting the broader trend toward digitalization in tax administration and corporate tax planning. Early foundational work in this area focused on establishing the theoretical compatibility between ML methodologies and the ALP, with researchers examining whether algorithmic approaches could satisfy existing legal and regulatory requirements for TP documentation [14]. These initial studies demonstrated that supervised learning techniques could effectively replicate and in many cases improve upon traditional comparable selection and pricing methodologies, while maintaining adherence to OECD guidelines when properly implemented and documented [15].

A significant stream of literature has examined the application of various PA techniques to the fundamental challenge of comparable company identification and selection, which represents a critical step in applying TNMM and other traditional TP methods. Research by Chen and colleagues demonstrated that ensemble methods combining RF and GBM achieved superior performance in identifying appropriate comparable companies compared to manual screening approaches, with particular improvements in handling high-dimensional financial and operational data [16]. This work highlighted the ability of ML algorithms to simultaneously consider multiple comparability factors including functional profile, asset intensity, risk profile, and market characteristics, thereby addressing one of the most subjective and contentious aspects of traditional TP practice. Subsequent research extended these findings by incorporating NLP techniques to analyze business descriptions and segment reporting data, enabling more nuanced functional comparability assessments [17].

The application of NN to TP analysis has generated considerable research interest, particularly regarding deep learning architectures capable of modeling complex value chain relationships and pricing dynamics. Studies have demonstrated that convolutional neural networks (CNN) and recurrent neural networks (RNN) can effectively capture temporal patterns in TP data, enabling more accurate forecasting of appropriate intercompany prices under varying market conditions [18]. However, this research also identified significant challenges related to model interpretability, as the black box nature of deep learning approaches conflicts with documentation requirements that necessitate clear explanations of pricing methodology [19]. This tension between predictive performance and regulatory acceptability has emerged as a central theme in the literature, with researchers exploring various approaches to explainable AI that can reconcile advanced modeling techniques with transparency requirements. Recent advances in knowledge-guided expert mixture architectures have demonstrated that domain-adapted large language models incorporating retrieval-augmented generation can achieve both high classification accuracy and interpretable outputs in tax analysis contexts, offering a promising approach for addressing similar challenges in transfer pricing applications [20].

Research examining regulatory perspectives on PA in TP contexts reveals substantial variation across jurisdictions in both the acceptance of algorithmic analyses and the standards applied to evaluate such methodologies. Comparative studies of tax authority guidance documents and audit practices indicate that while some jurisdictions have explicitly endorsed the use of advanced analytics subject to appropriate documentation standards, others maintain more conservative positions requiring primary reliance on traditional methods [21]. The BEPS Action 13 CbCR data has created new opportunities for tax authorities to deploy PA for risk assessment purposes, and research examining these applications demonstrates that predictive models can effectively identify high-risk TP arrangements warranting detailed examination [22]. However, concerns have been raised about potential biases in algorithmic risk scoring systems and the implications for taxpayer rights when automated systems drive audit selection decisions [23].

The economic substance analysis required under modern TP frameworks has also benefited from PA applications, with research demonstrating that ML techniques can enhance the identification and quantification of value drivers within complex global value chains. Studies employing classification algorithms and clustering techniques have shown promise in mapping functional contributions and risk allocations across multinational enterprises (MNEs), providing more systematic and data-driven support for profit allocation decisions [24]. This work addresses particularly challenging areas such as the valuation of intangible assets and the appropriate compensation for risk assumption, where traditional methodologies often rely heavily on subjective assessments [25]. Research has also examined how regression-based PA models can improve the estimation of arm's length returns by incorporating broader datasets and more sophisticated adjustment mechanisms for differences between controlled and uncontrolled transactions [26].

Literature addressing implementation challenges for PA in TP contexts identifies several critical success factors that determine whether organizations can effectively leverage these technologies. Data quality and availability emerge as primary concerns, with research demonstrating that the performance of ML models depends critically on access to comprehensive, accurate, and relevant financial and operational data spanning multiple years and jurisdictions [27]. Studies examining MNEs' experiences with PA implementation reveal that organizations often underestimate the data infrastructure requirements and the effort needed to integrate TP data with enterprise resource planning (ERP) systems and other corporate databases [28]. The importance of cross-functional collaboration between tax professionals, data scientists, and business units has been emphasized as essential for developing models that appropriately balance technical sophistication with practical applicability and regulatory defensibility [29].

Research on specific ML algorithms applied to TP problems has generated insights into the relative performance characteristics of different modeling approaches. Comparative studies evaluating support vector machines (SVM), decision trees, and ensemble methods for comparable selection tasks indicate that while ensemble approaches generally achieve superior predictive accuracy, simpler models may offer advantages in terms of interpretability and computational efficiency [30]. The application of unsupervised learning techniques including principal component analysis (PCA) and clustering algorithms has been explored for dimensionality reduction and pattern identification in complex TP datasets [31]. Research has also examined the potential of reinforcement learning approaches for dynamic TP optimization, though these applications remain largely theoretical due to regulatory constraints on prospective pricing optimization [32].

The regulatory implications of PA adoption extend beyond technical considerations to fundamental questions about the evolution of TP principles and administrative practices. Legal scholarship has examined whether existing regulatory frameworks adequately address the use of algorithmic decision-making in tax contexts, identifying potential gaps in areas such as algorithmic transparency requirements, standards for model validation and testing, and procedures for challenging automated determinations [33]. Research analyzing recent tax disputes involving PA-based TP analyses reveals emerging judicial perspectives on the evidentiary weight accorded to algorithmic studies and the standards applied in evaluating their reliability [34]. These cases highlight the importance of comprehensive documentation not only of modeling results but also of model development processes, including data sources, algorithm selection rationale, and validation procedures [35].

Studies examining the use of BDA in TP contexts have explored both opportunities and risks associated with incorporating increasingly granular transaction-level data into pricing analyses. Research demonstrates that access to detailed operational data can enable more precise comparable adjustments and more accurate profit allocations, particularly for complex value chains involving multiple jurisdictions and products [36]. However, concerns have been raised about potential privacy implications of extensive data collection and the challenges of managing cross-border data transfers in compliance with data protection regulations such as the European Union GDPR [37]. The intersection of TP compliance requirements and data privacy obligations represents an emerging area requiring further research and policy development [38].

Literature addressing the organizational change management aspects of PA implementation in TP functions reveals that successful adoption requires not only technical capabilities but also cultural shifts in how tax teams approach their work. Research examining change management practices identifies resistance to algorithmic decision-making as a significant barrier, particularly among experienced TP professionals who may view PA as threatening established expertise and professional judgment [39]. Studies highlight the importance of appropriate training programs that enable tax professionals to understand PA methodologies sufficiently to evaluate their appropriateness and interpret their results, even without developing deep technical expertise in data science [40]. The need for new roles bridging tax and analytics expertise has been identified, with research exploring optimal organizational structures for integrating these capabilities [41].

3 PREDICTIVE ANALYTICS METHODOLOGIES IN TRANSFER PRICING

The application of predictive analytics (PA) to transfer pricing (TP) encompasses a diverse array of methodologies, each offering distinct advantages for addressing specific analytical challenges within the TP process. Understanding the technical characteristics, appropriate applications, and limitations of these methodologies is essential for both practitioners seeking to implement these tools and regulators evaluating their use. Supervised learning algorithms represent the most widely adopted category of PA methodologies in TP applications, as they align naturally with the fundamental objective of predicting appropriate arm's length prices or profit margins based on historical data from comparable independent transactions [42]. These algorithms learn mapping functions from input features to output predictions through training on labeled datasets, where the labels represent known arm's length outcomes such as profit margins from independent companies or prices from uncontrolled transactions.

Random forests (RF) have emerged as particularly effective for comparable company selection and screening tasks within TP benchmarking studies. RF algorithms construct multiple decision trees during training and output the mode of the classes for classification tasks or mean prediction for regression tasks across individual trees [43]. The ensemble nature of RF provides several advantages for TP applications, including robustness to overfitting, ability to handle both numerical and categorical features without extensive preprocessing, and natural capacity to assess feature importance which aids in understanding which comparability factors most significantly influence outcomes. Research has demonstrated that RF models can effectively automate the initial screening of potential comparable companies by learning from historical selections made by TP experts, achieving classification accuracy rates exceeding 85 percent while significantly reducing the time required for comparable searches [44]. The interpretability of RF through feature importance scores also supports documentation requirements, as analysts can explain which characteristics drove inclusion or exclusion decisions for particular companies.

Gradient boosting machines (GBM), including popular implementations such as XGBoost and LightGBM, represent another powerful ensemble approach that has shown superior performance for regression tasks in TP contexts. GBM algorithms build models sequentially, with each new model attempting to correct errors made by the previous ensemble, resulting in highly accurate predictions for continuous outcomes such as profit margins or pricing levels [45]. The application of GBM to arm's length range determination has demonstrated particular promise, as these models can capture complex non-linear relationships between financial ratios, functional characteristics, and market conditions that influence appropriate comparable margins. Studies comparing GBM performance to traditional linear regression approaches for estimating arm's length ranges report improvements in both predictive accuracy and reduction in the width of predicted confidence intervals, suggesting more precise targeting of appropriate pricing levels [46].

Support vector machines (SVM) offer advantages for classification tasks in TP analysis, particularly when dealing with high-dimensional feature spaces and limited training data. SVM algorithms find optimal hyperplanes that maximize the margin between different classes in feature space, and through kernel tricks can efficiently handle non-linear decision boundaries [47]. In TP applications, SVM has been successfully employed for tasks such as classifying transactions into appropriate TP method categories, identifying high-risk pricing arrangements requiring detailed review, and predicting audit outcomes based on transaction characteristics. The mathematical rigor of SVM and its relatively transparent decision boundaries contribute to regulatory acceptability, though the choice of kernel function and hyperparameters requires careful validation to avoid overfitting [48].

Neural networks (NN), particularly deep learning architectures, represent the most sophisticated category of PA methodologies applied to TP, though their adoption faces significant challenges related to interpretability and regulatory acceptance. Deep NN with multiple hidden layers can learn hierarchical representations of data, potentially capturing subtle patterns in value creation and pricing dynamics that simpler models miss [49]. Recurrent neural networks (RNN) and long short-term memory (LSTM) networks have been explored for modeling temporal dependencies in TP data, such as how intercompany pricing should adjust in response to changing market conditions over multi-year periods. Convolutional neural networks (CNN) have been adapted for analyzing structured financial data and identifying patterns in multi-dimensional TP datasets. However, the black box nature of deep learning models creates substantial documentation challenges, as explaining precisely how a deep NN arrived at a particular pricing recommendation may be impossible even for the data scientists who built the model [50]. This opacity conflicts fundamentally with TP regulations requiring clear articulation of pricing methodology and economic reasoning.

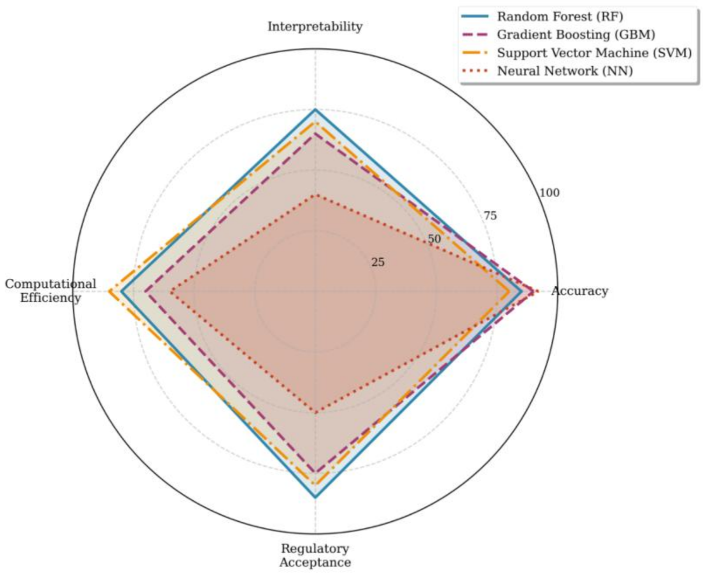


Figure 1 Comparison of ML Algorithm Performance Across Accuracy, Interpretability, Computational Efficiency, and Regulatory Acceptance for Transfer Pricing Applications

To address the interpretability challenge while retaining predictive power, researchers and practitioners have increasingly adopted explainable AI (XAI) techniques in TP applications. SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) represent two prominent XAI approaches that can generate explanations for predictions made by complex ML models [51]. SHAP values, grounded in cooperative game theory, quantify the contribution of each input feature to a particular prediction, enabling TP practitioners to understand and document which factors drove a specific pricing recommendation. LIME generates locally faithful explanations by fitting interpretable models around individual predictions, providing insight into model behavior for specific transactions even when the global model is highly complex. The integration of XAI techniques with advanced PA models represents a promising path toward reconciling predictive performance with regulatory transparency requirements [52].

Figure 1 compares four primary ML algorithms across key performance dimensions for TP applications. Random forests and gradient boosting machines achieve high accuracy while maintaining moderate interpretability through feature importance scores, making them well-suited for comparable selection tasks. Support vector machines offer strong regulatory acceptance due to transparent decision boundaries but show lower accuracy for complex pricing analyses. Neural networks demonstrate superior accuracy for intricate value chain modeling but score lowest on interpretability and regulatory acceptance due to their black-box nature. This trade-off analysis informs algorithm selection decisions, emphasizing that optimal choices depend on specific application requirements and jurisdictional regulatory expectations.

Unsupervised learning methodologies play complementary roles in TP analytics, particularly for exploratory analysis and pattern discovery in situations where labeled training data is limited or unavailable. Clustering algorithms such as k-means, hierarchical clustering, and density-based spatial clustering of applications with noise (DBSCAN) can identify natural groupings within populations of potential comparable companies or transactions, helping analysts understand the structure of available data and identify potential peers that might not emerge from traditional screening criteria [53]. Principal component analysis (PCA) and other dimensionality reduction techniques enable visualization and exploration of high-dimensional TP datasets, helping analysts identify which combinations of features best distinguish between different groups of comparables or explain variation in arm's length outcomes. These unsupervised approaches often serve as valuable preprocessing steps that enhance the performance and interpretability of subsequent supervised learning models.

4 REGULATORY IMPLICATIONS AND COMPLIANCE CHALLENGES

The integration of predictive analytics (PA) into transfer pricing (TP) practices creates profound regulatory implications that extend across multiple dimensions of tax administration, compliance, and policy development. Tax authorities worldwide face the dual challenge of evaluating PA-based analyses submitted by multinational enterprises (MNEs) while simultaneously deploying similar technologies for their own enforcement and compliance activities. This parallel adoption creates a dynamic regulatory environment characterized by evolving standards, jurisdictional variations, and ongoing debates about the appropriate role of algorithmic decision-making in tax determination [54]. Understanding these regulatory implications is essential for MNEs seeking to leverage PA effectively while managing compliance risks and maintaining defensible positions.

The fundamental question of whether PA-based analyses satisfy the arm's length principle (ALP) and comply with Organisation for Economic Co-operation and Development (OECD) Transfer Pricing Guidelines has been addressed differently across jurisdictions. Some tax authorities have issued guidance explicitly acknowledging that advanced statistical methods and machine learning (ML) algorithms may be acceptable for certain aspects of TP analysis, provided that the methodologies are properly documented, validated, and applied in a manner consistent with established TP principles [55]. These jurisdictions typically require that PA applications supplement rather than replace traditional analyses, with algorithmic results subject to expert review and adjustment based on qualitative factors not captured in models. Other jurisdictions have maintained more conservative positions, expressing concerns about the transparency and auditability of complex algorithms and requiring primary reliance on conventional methodologies with PA serving only as supporting evidence [56].

Documentation requirements represent a critical compliance challenge for MNEs employing PA in TP contexts, as standard documentation practices developed for traditional analyses may not adequately address the unique characteristics of algorithmic approaches. Tax authorities generally expect documentation to explain not only the results of PA models but also the model development process, including data sources and quality, feature selection rationale, algorithm choice justification, training and validation procedures, and testing for potential biases or errors [57]. This level of detail requires close collaboration between tax and data science teams and may necessitate disclosure of technical information that organizations consider proprietary or commercially sensitive. The challenge is particularly acute for MNEs using proprietary or licensed PA tools, where full transparency regarding algorithmic implementation may be limited by vendor restrictions [58].

The evidentiary standards applied to PA-based analyses in tax disputes and litigation have begun to emerge through case law and administrative proceedings, though this body of precedent remains limited. Early decisions suggest that courts and tribunals are generally willing to consider algorithmic analyses as evidence, but apply rigorous standards regarding the quality of data inputs, appropriateness of methodology for the specific application, and qualifications of experts who developed and interpreted the models [59]. Cases have emphasized the importance of independent validation of PA models, with particular scrutiny applied to prevent overfitting or other forms of model bias that could generate misleading results. The burden of proof considerations in TP disputes may be affected by PA adoption, as taxpayers employing sophisticated analytical methods may face heightened expectations regarding the rigor and comprehensiveness of their supporting evidence [60].

Data privacy and cross-border data transfer regulations create additional compliance complexity for MNEs seeking to implement PA for TP purposes, particularly for organizations operating across multiple jurisdictions with varying data protection requirements. The application of PA typically requires aggregating and analyzing transaction-level data from multiple entities and jurisdictions, which may involve transfer of personal data subject to restrictions under regulations such as the European Union General Data Protection Regulation (GDPR) and similar frameworks in other regions [61]. MNEs must ensure that their PA systems comply with data localization requirements, obtain necessary consents for data processing, and implement appropriate technical and organizational measures to protect data security. The tension between TP compliance requirements demanding comprehensive data analysis and data privacy regulations limiting data collection and transfer represents an ongoing challenge requiring careful legal and technical navigation [62].

Table 1 Regulatory Positions on Predictive Analytics in Transfer Pricing across Major Tax Jurisdictions showing Acceptance Levels, Documentation Requirements, and Key Restrictions

Jurisdiction	Key Guidance Document	Acceptance Level	Documentation Requirement	Key Restrictions/Notes
United States	IRS Revenue Procedure 2015-41	Medium	Moderate	Requires economic substance analysis
United Kingdom	HMRC Transfer Pricing Guidelines	High	Detailed	Explicitly endorses advanced analytics
Germany	BMF Administrative Principles 2	Medium	High	Primary reliance on traditional methods
France	BOI-BIC-BASE-80-10-20	Medium-High	Moderate	Accepts with proper documentation
China	SAT Bulletin [2017] No.6	High	Very High	Strong focus on CbCR data analytics
Singapore	IRAS e-Tax Guide (5th Edition)	High	Moderate	Encourages innovative approaches
Australia	ATO PCG 2019/1	Medium	High	Requires transparency and validation

Tax authority adoption of PA for audit selection and risk assessment creates a parallel set of implications for MNE compliance strategies. Many tax administrations have deployed predictive models to analyze country-by-country reporting (CbCR) data and other information returns, identifying taxpayers with TP arrangements that warrant detailed examination [63]. These risk assessment systems typically employ ML algorithms to identify patterns associated with aggressive TP planning, such as high-value intangible transfers to low-tax jurisdictions, profit allocations inconsistent with functional analysis, or outlier financial results compared to industry norms. While these systems can enhance the efficiency and effectiveness of tax administration by directing resources toward highest-risk cases, they also raise concerns about algorithmic bias, lack of transparency in audit selection decisions, and potential for false positives that subject compliant taxpayers to burdensome audits [64].

The advance pricing agreement (APA) process represents an area where PA both offers significant opportunities and presents unique challenges. PA models can support APA applications by providing more robust forecasts of appropriate pricing ranges over multi-year periods, potentially reducing uncertainty and controversy [65]. However, tax authorities evaluating APA requests based on PA may require extensive information about model construction, assumptions, and sensitivity analysis to assess the reliability of forecasts. The ongoing maintenance and updating of PA models throughout APA terms creates additional considerations, as models may require recalibration in response to changing business conditions or data availability. Questions about what level of deviation from PA-based forecasts would constitute APA non-compliance, and how such deviations should be addressed, represent emerging issues in APA practice.

Table 1 synthesizes regulatory stances toward PA in TP across seven major jurisdictions. The United Kingdom demonstrates the most progressive position with explicit HMRC endorsement, while Germany maintains the most conservative approach requiring primary reliance on traditional methods. Documentation requirements vary significantly, with China and Australia mandating extensive technical specifications including algorithm validation procedures. These jurisdictional variations create substantial compliance challenges for MNEs operating across multiple tax regimes, necessitating tailored documentation strategies. The trend toward conditional acceptance—requiring PA to supplement rather than replace traditional analyses—reflects regulators balancing innovation encouragement with established ALP interpretation principles.

5 IMPLEMENTATION CONSIDERATIONS AND BEST PRACTICES

Successful implementation of predictive analytics (PA) in transfer pricing (TP) requires careful attention to technical, organizational, and strategic considerations that extend beyond algorithm selection and model development. Organizations that have effectively integrated PA into TP functions typically follow systematic approaches addressing data infrastructure, governance frameworks, capability development, and stakeholder engagement. The foundation of any PA implementation is adequate data infrastructure capable of collecting, storing, and processing the diverse financial and operational data required for TP analysis. MNEs must assess whether their existing enterprise resource planning (ERP) systems and data warehouses can support PA requirements, including transaction-level detail, multi-year historical data, and integration of both internal financial data and external market information [66]. Data quality initiatives focusing on completeness, accuracy, consistency, and timeliness are essential prerequisites, as PA model performance degrades significantly when trained on incomplete or erroneous data.

Governance frameworks for PA in TP should address key questions about model ownership, validation requirements, update frequencies, and decision-making authority. Leading organizations typically establish cross-functional governance committees including representatives from tax, finance, IT, and data analytics functions to oversee PA deployments [67]. These committees define standards for model documentation, establish protocols for model validation and testing, determine circumstances under which PA results require expert review before application, and manage the balance between algorithmic recommendations and professional judgment. Clear governance helps ensure that PA tools are applied appropriately and consistently while preventing over-reliance on models without adequate human oversight.

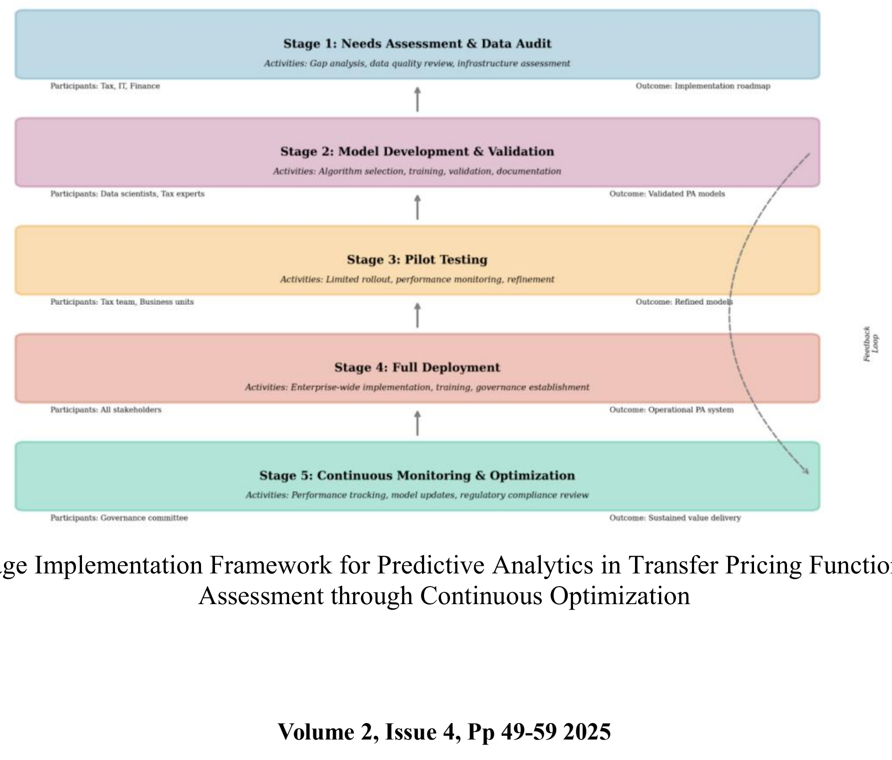


Figure 2 Five-Stage Implementation Framework for Predictive Analytics in Transfer Pricing Functions showing needs Assessment through Continuous Optimization

Capability development represents a critical success factor, as effective PA implementation requires building new skills within tax teams and fostering productive collaboration between tax and analytics professionals. MNEs approach capability building through various strategies including hiring data scientists with specialized training in tax applications, upskilling existing tax professionals through targeted analytics training programs, or establishing centers of excellence combining deep tax and analytics expertise [68]. The optimal approach depends on organizational size, resource availability, and strategic importance of PA to overall TP risk management. Regardless of structure, successful implementations emphasize the importance of tax professionals developing sufficient analytical literacy to evaluate PA outputs critically and understand their limitations, even without becoming expert data scientists themselves.

Figure 2 presents a five-stage framework for PA implementation in TP functions. Stage 1 involves needs assessment and data audit to identify high-value applications. Stage 2 focuses on model development with close collaboration between data scientists and TP experts to ensure regulatory defensibility. Stage 3 conducts pilot testing against traditional methods. Stage 4 executes enterprise-wide deployment with comprehensive training and governance frameworks. Stage 5 establishes continuous monitoring and optimization. The feedback loop connecting Stage 5 back to Stage 2 reflects the iterative nature of PA systems requiring periodic recalibration as business conditions and regulatory standards evolve.

Change management and stakeholder engagement processes help address resistance to PA adoption and ensure that new analytical capabilities are effectively integrated into existing workflows. Experienced TP professionals may initially view PA as threatening established expertise or may be skeptical about the reliability of algorithmic approaches compared to traditional judgment-based methods [69]. Addressing these concerns requires transparent communication about the role of PA as enhancing rather than replacing professional expertise, demonstration of PA value through pilot projects showing concrete improvements in accuracy or efficiency, and involvement of senior TP leaders as champions for analytics adoption. Engagement with external stakeholders including auditors and tax authorities helps ensure that PA implementations will be accepted in compliance and dispute contexts.

6 CONCLUSION

The application of predictive analytics to transfer pricing represents a transformative development with significant potential to enhance accuracy, efficiency, and defensibility of intercompany pricing practices. This review has demonstrated that diverse PA methodologies including random forests, gradient boosting machines, support vector machines, and neural networks offer substantial capabilities for addressing traditional challenges in comparable selection, arm's length range determination, and economic analysis. These technologies enable processing of larger datasets, identification of more nuanced patterns, and more systematic approaches to aspects of TP analysis that have historically relied heavily on subjective judgment[70]. The documented performance improvements in various TP applications suggest that PA will become increasingly integral to how sophisticated MNEs approach TP compliance and how tax authorities conduct risk assessment and audit activities.

However, successful PA adoption requires navigating significant regulatory, technical, and organizational challenges. The regulatory landscape remains fragmented, with varying levels of acceptance and divergent documentation expectations across jurisdictions. Tax authorities and policymakers must develop clearer guidance regarding acceptable PA applications, appropriate transparency standards, and evidentiary requirements that balance innovation with established legal principles[71]. The interpretability challenge inherent in sophisticated ML models demands continued development and adoption of explainable AI techniques that can reconcile predictive power with regulatory transparency requirements. Documentation practices must evolve to adequately address the unique characteristics of algorithmic analyses while remaining practical for resource-constrained tax functions.

The parallel adoption of PA by both taxpayers and tax authorities creates a dynamic environment where technological capabilities on both sides of the compliance relationship are rapidly evolving. This development offers opportunities for more efficient administration and reduced compliance costs, but also raises important questions about algorithmic fairness, procedural justice, and appropriate safeguards against potential biases in automated decision-making systems. The intersection of TP compliance requirements with data privacy regulations presents ongoing challenges requiring coordination between tax and legal functions[72]. Organizations must carefully design PA implementations to respect data protection principles while meeting analytical requirements.

Implementation success depends critically on adequate data infrastructure, robust governance frameworks, appropriate capability development, and effective change management. MNEs should approach PA adoption strategically, beginning with pilot applications in well-defined use cases, building internal expertise through training and hiring, and developing documentation practices that will withstand regulatory scrutiny. Cross-functional collaboration between tax, finance, IT, and data analytics teams is essential for developing solutions that balance technical sophistication with practical applicability and regulatory defensibility.

Looking forward, continued research is needed in several areas to support responsible PA adoption in TP contexts. Development of industry-specific benchmarks and validation standards would provide useful reference points for assessing model quality and appropriateness. Further empirical studies examining the accuracy of PA approaches compared to traditional methods across diverse transaction types and industries would build evidence supporting broader adoption. Research addressing algorithmic fairness and bias detection in TP applications could help ensure that PA deployment promotes rather than undermines equitable international taxation. Exploration of emerging technologies including blockchain for real-time TP documentation and quantum computing for complex optimization problems may reveal additional opportunities for innovation.

The transformation of TP practice through PA is inevitable given the clear performance advantages these technologies offer and their increasing adoption by both taxpayers and authorities. The challenge for stakeholders is to guide this transformation in directions that preserve fundamental tax principles while enabling beneficial innovation. With appropriate attention to regulatory frameworks, technical standards, and implementation practices, PA can enhance the effectiveness and fairness of the international TP system for all participants.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Mwape V, Mwange A, Hamalengwa M. Comparative jurisprudence on the application of the arm’s length principle: A global survey of transfer pricing regulation and its implications for Zambia. African Journal of Commercial Studies, 2025, 6(3): 19-26.
[2] Pichhadze A. 'Transfer pricing and the arm’s length principle after BEPS': Book review. British Tax Review, 2020(2).
[3] Khare A, Ojala A, Baber W W. Sustainable international business models in a digitally transforming world. Routledge, 2022.
[4] Yang S, Ding G, Chen Z, et al. GART: Graph Neural Network-based Adaptive and Robust Task Scheduler for Heterogeneous Distributed Computing. IEEE Access, 2025.
[5] Safaei R B. Transfer pricing rules for intangibles: Implementation and practical challenges. In: Transfer Pricing Rules and Tax Audits: Effects on Investments & Compliance, 2023: 48.
[6] Castagnari F. Arbitration tax proceedings for international tax-treaty related disputes, 2023.
[7] Blum D W. Normativity in international tax law: A legal theoretical inquiry into contemporary international tax discourse, 2023.
[8] Seidel S, Watson R T. Integrating explanatory/predictive and prescriptive science in information systems research. Communications of the Association for Information Systems, 2020, 47(1): 49.
[9] Mashiri E, Dzomira S, Canicio D. Transfer pricing auditing and tax forestalling by multinational corporations: A game theoretic approach. Cogent Business & Management, 2021, 8(1): 1907012.
[10] Montesinos López O A, Montesinos López A, Crossa J. Fundamentals of artificial neural networks and deep learning. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction. Cham: Springer International, 2022: 379-425.
[11] AL-Mashaqbeh K M. The relationship between e-government and job performance. Daralkhalij for Publishing and Distribution, 2023.
[12] Edwards B D, Travis III L F. Introduction to criminal justice. Routledge, 2019.
[13] Mercurio B, Yu R. An AI policy for the (near) future. Addressing Impediments to Digital Trade, 2021: 73-104.
[14] Lee Y S. Law and development: Theory and practice. Routledge, 2022.
[15] Chyz J A, Gaertner F B, Kausar A, et al. Overconfidence and corporate tax policy. Review of Accounting Studies, 2019, 24(3): 1114-1145.
[16] Chen M C, Li H Y. Are Taiwanese descendant family firms more aggressive in tax avoidance?. Asia-Pacific Journal of Accounting & Economics, 2025: 1-18.
[17] Thayyib P V. Tax avoidance strategies under radar: Policy considerations on US multinational corporations operating in emerging markets. Cogent Business & Management, 2025, 12(1): 2507236.
[18] Hanlon M. Tax research in accounting: A look forward. The Accounting Review, 2025, 100(6): 359-371.
[19] Dyreng S D, Hanlon M, Maydew E L. When does tax avoidance result in tax uncertainty?. The Accounting Review, 2019, 94(2): 179-203.
[20] Tan Y, Wu B, Cao J, et al. LLaMA-UTP: Knowledge-guided expert mixture for analyzing uncertain tax positions. IEEE Access, 2025.
[21] Aisyah S. Influence of taxes on capital flows and foreign direct investment. Golden Ratio of Taxation Studies, 2025, 5(1): 17-28.
[22] De Simone L, Olbert M. Real effects of private country-by-country disclosure. The Accounting Review, 2022, 97(6): 201-231.
[23] Favaretto M, De Clercq E, Elger B S. Big data and discrimination: Perils, promises and solutions—A systematic review. Journal of Big Data, 2019, 6(1): 1-27.
[24] Jan H A, Bussmann R W, Mir T A, et al. Cross-regional ethnomedicinal study of plants-based herbal tea in six tehsils of district Buner, northern Pakistan. Pakistan Journal of Botany, 2024, 56(5).
[25] Ali S. Corporate taxation and intangible assets: A systematic literature review and future research trends. Eurasian Business and Economics Perspectives, 2022: 171-189.
[26] Shehaj P, Weichenrieder A J. Corporate income tax, IP boxes and the location of R&D. International Tax and Public Finance, 2024, 31(1): 203-242.

[27] Nepivoda L. Profit shifting of multinational enterprises: A refined FDI approach, 2019.
[28] Petr J. Profit shifting of multinational corporations worldwide. arXiv preprint arXiv:2201.08444, 2022.
[29] Drake K D, Goldman N C, Murphy F. Foreign employment, income shifting, and tax uncertainty. The Accounting Review, 2022, 97(2): 183-212.
[30] Tian L, Yang X, Qin L. Shareholder power and tax avoidance: Evidence from geographic concentration of institutions. Applied Economics, 2025: 1-15.
[31] Singh S, Ramkumar K R, Kukkar A. Machine learning techniques and implementation of different ML algorithms. In: 2021 2nd Global Conference for Advancement in Technology (GCAT). IEEE, 2021: 1-6.
[32] Katambara Z. Multivariate analysis of evaporation drivers in Mbeya, Tanzania using principal component analysis, 2024.
[33] Goldwaser A, Thielscher M. Deep reinforcement learning for general game playing. In: Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(02): 1701-1708.
[34] Mukherjee A, Chang H H. Agentic AI: Autonomy, accountability, and the algorithmic society. arXiv preprint arXiv:2502.00289, 2025.
[35] Blouin J, Robinson L A. Accounting for the profits of multinational enterprises: Double counting and misattribution of foreign affiliate income. SSRN, 2023.
[36] Richtárová M A. Global distribution of revenue loss from corporate tax avoidance, 2022.
[37] Elumilade O O, Ogundeji I A, Achumie G O, et al. Optimizing corporate tax strategies and transfer pricing policies to improve financial efficiency and compliance. Journal of Advance Multidisciplinary Research, 2022, 1(2): 28-38.
[38] Παπαδόπουλος A. The implementation of the General Data Protection Regulation (GDPR) in the EU and Greece: Procedures, risks, challenges and impacts in the context of good governance.
[39] Marelli L, Testa G, Van Hoyweghen I. Big Tech platforms in health research: Re-purposing big data governance in light of the General Data Protection Regulation’s research exemption. Big Data & Society, 2021, 8(1): 20539517211018783.
[40] Wang Y, Ding G, Zeng Z, et al. Causal-aware multimodal transformer for supply chain demand forecasting: Integrating text, time series, and satellite imagery. IEEE Access, 2025.
[41] Lubis C W, Kesuma S A, Muda I. Corporate governance as a determiner of tax avoidance: A literature review study. International Journal of Social Service and Research, 2023, 3(3): 690-696.
[42] Chen S, Schuchard K, Stomberg B. Media coverage of corporate taxes. The Accounting Review, 2019, 94(5): 83-116.
[43] Fantin Irudaya Raj E, Balaji M. Application of deep learning and machine learning in pattern recognition. In: Advance Concepts of Image Processing and Pattern Recognition: Effective Solution for Global Challenges. Singapore: Springer Singapore, 2022: 63-89.
[44] Shahid M N, Ullah S, Jamal A, et al. In silico prediction of drug based potential inhibitors against coat protein of tobacco ringspot virus. Pakistan Journal of Botany, 2024, 56(5).
[45] Schonlau M, Zou R Y. The random forest algorithm for statistical learning. The Stata Journal, 2020, 20(1): 3-29.
[46] Morán-Fernández L, Bolón-Canedo V, Alonso-Betanzos A. Do we need hundreds of classifiers or a good feature selection?. In: ESANN, 2020: 399-404.
[47] Zhang P, Jia Y, Shang Y. Research and application of XGBoost in imbalanced data. International Journal of Distributed Sensor Networks, 2022, 18(6): 15501329221106935.
[48] Rufo D D, Debelee T G, Ibenthal A, Negera W G. Diagnosis of diabetes mellitus using gradient boosting machine (LightGBM). Diagnostics, 2021, 11(9): 1714.
[49] Ghosh S, Dasgupta A, Swetapadma A. A study on support vector machine based linear and non-linear pattern classification. In: 2019 International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2019: 24-28.
[50] Gao H, Fam P S, Low H C. A comparative study between support vector machine and support vector data description in bank telemarketing. In: Proceedings 63rd ISI World Statistics Congress, 2021, 11: 16.
[51] Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[52] Abou El Houda Z, Brik B, Khoukhi L. "Why should I trust your IDS?": An explainable deep learning framework for intrusion detection systems in Internet of Things networks. IEEE Open Journal of the Communications Society, 2022, 3: 1164-1176.
[53] Chen J, Jordan M. LS-Tree: Model interpretation when the data are linguistic. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(04): 3454-3461.
[54] Molnar C. Interpretable machine learning. Lulu.com, 2020.
[55] Mai N T, Fang Q, Cao W. Measuring Student Trust and Over-Reliance on AI Tutors: Implications for STEM Learning Outcomes. International Journal of Social Sciences and English Literature, 2025, 9(12): 11-17.
[56] Valderrama I J M. BEPS principal purpose test and customary international law. Leiden Journal of International Law, 2020, 33(3): 745-766.
[57] Hezarkhani B, Demirel G. Reflections of Global Supply Chains in Mirror Trade Data. SSRN 5391771, 2025.
[58] Kraievskyi V, Muravskyi O. Tax compliance control and audit of transfer pricing transactions. Економіка розвитку систем, 2024, 6(1): 37-45.
[59] Petruzzi R, Lang M. Transfer Pricing Developments around the world 2023, Vol. 48. Kluwer Law International BV, 2022.

[60] Screpante M. The OECD transfer pricing guidelines and value creation: income allocation or anti-avoidance tool? 2023.
[61] Danish S, Rehman R A, Maqbool A, et al. Effects of benzyl amino purine on growth, antioxidants and chlorophyll contents of Phaseolus vulgaris under heat stress. Pakistan Journal of Botany, 2024, 56(5).
[62] Panayi C H. The relationship between EU and international tax law. In: Research Handbook on European Union Taxation Law. Edward Elgar Publishing, 2020: 119-141.
[63] Fjord L B, Koerver Schmidt P. The Digital Transformation of Tax Systems-Progress, Pitfalls and Protection in a Danish Context. Indiana Journal of Global Legal Studies, 2022: 22-10.
[64] Voigt P, von dem Bussche A. Scope of application of the GDPR. In: The EU general data protection regulation (GDPR) a practical guide. Springer Nature Switzerland, 2024: 9-36.
[65] Kühling J, Klar M, Sackmann F, Sackmann K K. Datenschutzrecht. CF Müller, 2025.
[66] Framework R. OECD Online Dispute Resolution Framework, 2024.
[67] Kheya T A, Bouadjenek M R, Aryal S. The pursuit of fairness in artificial intelligence models: A survey. arXiv:2403.17333, 2024.
[68] Kofler G, Lang M, Pistone P, et al. CJEU-Recent Developments in Direct Taxation 2023. Series on International Tax Law, Vol.142, 2024.
[69] Lin H, Liu W. Symmetry-Aware Causal-Inference-Driven Web Performance Modeling: A Structure-Aware Framework for Predictive Analysis and Actionable Optimization. Symmetry, 2025, 17(12): 2058.
[70] Kauppila E. The Role of Enterprise Architecture in Strategy Work, 2023.
[71] Sawang S, Kivits R A. Revolutionizing family businesses with artificial intelligence: a perspective article. Journal of Family Business Management, 2024, 14(4): 802-807.
[72] Schellenberger K W. Meta-Analysis Examining Transformational Leadership and Successful Organizational Change. Doctoral dissertation, Keiser University, 2019.