Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/AU04.20250010
Published: November 25, 2025

Reinforcement Learning for Financial Resilience: Optimizing Supply Chain Liquidity Under Uncertainty

Yehan Dou and Peter Shi
Department of Management
Macquarie University
Sydney, Australia
yehan.dou@mq.edu.au, peter.shi@mq.edu.au

Ruoyu Li
School of Engineering and Information Science
The University of Wollongong
Wollongong, Australia
lruoyu@uow.edu.au

Abstract

In the evolving landscape of global trade and digital finance, optimizing liquidity and managing credit risk in supply chain finance (SCF) have become pressing challenges. This study proposes a novel reinforcement learning (RL)-based framework designed to optimize cash flow strategies within SCF ecosystems dynamically. By integrating machine learning-driven credit evaluation, multi-agent decision modeling, and adaptive policy learning, the framework addresses long-standing inefficiencies in conventional SCF practices, including rigid pricing structures, static credit scoring methodologies, and misaligned stakeholder incentives. The system formulates SCF interactions as a stochastic Markov Decision Process (MDP) and incorporates game-theoretic mechanisms to capture negotiation dynamics between borrowers and lenders. Numerical experiments, conducted using both synthetic data and empirical records from multinational corporations, demonstrate that the RL-based policy significantly reduces borrower financing costs, enhances default risk management, and improves platform profitability across a range of macro-financial conditions. The proposed approach provides a scalable, interpretable, and resilient decision-support system, underscoring the transformative potential of artificial intelligence in shaping the future of SCF.

Keywords
Reinforcement learning, Supply chain finance, Optimization, Digital finance

1. Introduction
In today's volatile and interconnected global economy, firms face growing pressure to maintain liquidity and ensure agile working capital management. Disruptive macroeconomic forces—including geopolitical tensions, inflationary shocks, and trade decoupling—have heightened the vulnerability of supply networks, especially for small and medium-sized enterprises (SMEs) embedded in complex global value chains (Luo & Witt, 2021). These developments have elevated cash flow optimization from a tactical concern to a strategic imperative for business resilience and financial stability.

In response to these challenges, supply chain finance (SCF) has emerged as a key financial innovation aimed at bridging liquidity gaps between upstream suppliers and downstream buyers. Despite the growing adoption of digital financial platforms and tools such as AI-powered lending, decentralized finance (DeFi), and mobile banking infrastructure (Ronchini et al., 2024), most existing SCF models remain limited by rule-based heuristics, rigid credit scoring systems, and linear optimization frameworks. These static approaches fail to account for real-time market dynamics, behavioral feedback, and stakeholder interdependence, thereby constraining their effectiveness in volatile environments.

Against this backdrop, reinforcement learning (RL) offers a powerful paradigm for modelling and solving high-dimensional, sequential decision problems under uncertainty. RL agents learn optimal policies through continuous interaction with the environment and maximize cumulative rewards—a structure well-aligned with dynamic credit risk adjustment and adaptive liquidity allocation (Sutton & Barto, 2018; Li et al., 2019). While RL has seen increasing success in financial domains such as portfolio management, option pricing, and algorithmic trading (Deng et al., 2016; Jiang et al., 2017; Fischer, 2018), its application in SCF remains scarce, primarily due to the complexity of modelling heterogeneous agents, logistical constraints, asymmetric information, and multi-agent coordination.

This study addresses this research gap by proposing a novel reinforcement learning-based framework for optimizing cash flow strategies within SCF platforms. The framework integrates machine learning-based credit evaluation, multi-agent game-theoretic modelling, and adaptive policy learning to capture the complexity of real-world SCF ecosystems. SCF dynamics are formulated as a stochastic Markov Decision Process (MDP), and the model incorporates both bilateral lender-borrower interactions and network-level liquidity allocation strategies.

To validate the proposed framework, we conduct numerical experiments using both synthetic simulation environments and empirical data from multinational corporations. The experiments benchmark the RL-based policy against static and rule-based alternatives across key metrics such as financing cost, credit default risk, and platform profitability. Results demonstrate that the RL framework reduces borrower financing costs by 15%, improves default risk mitigation, and enhances system-wide liquidity allocation under both normal and stress scenarios.

This study makes several key contributions. Theoretically, it advances the literature by embedding adaptive, multi-agent learning architectures into SCF systems and highlighting the interplay between micro-level credit risk and macro-level liquidity distribution. Methodologically, it develops a robust, scalable, and explainable decision framework that integrates predictive analytics, dynamic reward shaping, and inter-agent coordination. Practically, it offers actionable insights for designing resilient and inclusive digital financial infrastructures that can support SME financing, crisis response, and credit accessibility in supply chain ecosystems.

1.1 Objectives

The primary objective of this study is to develop an RL-based decision framework that addresses the limitations of existing SCF models in managing liquidity and credit risk under uncertainty. Specifically, the study aims to formulate SCF interactions as a stochastic Markov Decision Process that incorporates multi-agent dynamics, heterogeneous firm behaviors, and macro-financial volatility. It seeks to integrate machine learning-based credit evaluation and game-theoretic coordination into a unified, adaptive policy optimization architecture. Furthermore, the research aims to empirically evaluate the proposed framework using both synthetic simulations and real-world data from SMEs supply networks, with a focus on financing cost reduction, credit risk mitigation, and system-level profitability. The final objective is to enhance policy interpretability through explainable AI techniques and provide actionable insights for SCF platform design and financial innovation.

2. Literature Review

2.1 Supply chain finance

SCF has become a strategic lever for enhancing liquidity and financial resilience across global production networks. Classic SCF instruments such as factoring, reverse factoring, and dynamic discounting are designed to ease working capital constraints—particularly for SMEs—by bridging cash flow gaps between buyers and suppliers (Gelsomino et al., 2016). However, these mechanisms often rely on rigid credit evaluation systems, typically based on static financial ratios and retrospective data. This creates significant barriers for SMEs with limited financial disclosures and renders such systems incapable of responding to real-time operational signals or market volatility (Kouvelis et al., 2020). These limitations underscore the need for more adaptive, data-driven credit scoring mechanisms, which our study addresses through the integration of machine learning-based credit evaluation into an RL framework.

Another major issue lies in the deterministic nature of traditional SCF optimization models, which assume stable environments and fail under abrupt shocks such as geopolitical crises, supply chain disruptions, or liquidity crunches. In volatile markets, static pricing and financing policies become ineffective or even counterproductive. Although some studies have attempted to incorporate big data analytics for early risk warning (e.g., Gong et al., 2022), few offer real-time, sequential decision-making frameworks capable of adjusting financing policies dynamically. This represents a key research gap, which this paper fills by modelling SCF as a stochastic Markov Decision Process (MDP) and deploying RL agents to optimize financing strategies under uncertainty.

A further challenge in SCF lies in the misalignment of incentives across multiple stakeholders—banks aim to maximize returns, suppliers seek immediate liquidity, and buyers focus on cost minimization. These competing objectives often lead to inefficient financing outcomes and hinder collaborative decision-making (Pfohl & Gomm, 2009). While digital platforms and fintech solutions have improved transparency, they rarely redesign the underlying logic of financing policy coordination. Most lack mechanisms for strategic bargaining or modelling multi-agent interactions. In contrast, our proposed RL-based framework embeds multi-agent game-theoretic coordination, allowing agents to learn cooperative or adversarial strategies over time, thus aligning financing policies with both firm-level and system-wide objectives.

2.2 Reinforcement Learning in Financial Decision-Making and SCF Applications

RL has emerged as a promising approach in financial decision-making due to its capacity to model sequential, uncertain, and interactive environments. Unlike supervised learning, RL agents optimize decisions by learning through trial-and-error interactions with the environment, guided by cumulative rewards (Li et al., 2019). This paradigm aligns naturally with financial tasks that involve intertemporal trade-offs, such as investment allocation, credit decisions, and risk-adjusted pricing. By balancing exploration and exploitation, RL is particularly well-suited to dynamic, partially observable decision contexts.

In capital markets, RL has been successfully applied to a range of problems, including portfolio optimization (Jiang et al., 2017), option pricing (Deng et al., 2016), and algorithmic trading (Fischer, 2018). These studies demonstrate RL's potential to adapt to shifting market conditions, respond to delayed rewards, and outperform static or rule-based strategies. More recent applications have extended to corporate finance and credit risk modelling, where sequential decisions and feedback loops play a key role (Moody & Saffell, 2001). However, these successes are largely confined to well-structured, single-agent environments with clear feedback mechanisms.

By contrast, the application of RL in SCF remains limited and fragmented. SCF ecosystems are characterized by heterogeneous agents (suppliers, buyers, financiers), non-financial dependencies (e.g., logistics constraints), and exposure to exogenous shocks (e.g., geopolitical risk, inflation). Only a handful of studies have explored RL in this context—for example, Wang et al. (2023) applied Q-learning to optimize invoice discounting, and Ronchini et al. (2024) reviewed the role of AI in SCF digitalization. However, most existing models address narrow decision tasks and rely on over-simplified assumptions, such as full observability or centralized control. They often fail to capture inter-agent negotiations, dynamic pricing, and risk-sharing mechanisms, which are central to SCF operations.

Moreover, current RL applications in SCF often ignore multi-agent dynamics and strategic interactions. In real-world SCF systems, financing decisions involve bargaining between lenders and borrowers, coordination across the supply network, and trade-offs between systemic liquidity and firm-level profitability. These complexities require advanced techniques such as multi-agent reinforcement learning (MARL) or game-theoretic RL formulations. Yet, few studies incorporate these perspectives, leaving a critical gap in modelling realistic and scalable policy architectures for SCF platforms.

In addition to theoretical limitations, practical deployment of RL in SCF faces challenges related to data availability, model interpretability, and regulatory compliance. While recent advances in explainable AI and policy distillation offer potential solutions (Bastani et al., 2022), their integration into operational financial systems remains underexplored. As such, there is an urgent need for intelligent SCF frameworks that embed adaptive policy learning, support multi-agent coordination, and remain robust under uncertainty. This study addresses this gap by proposing a novel RL-based framework that combines predictive analytics, credit scoring, and game-theoretic learning to optimize SCF outcomes for all stakeholders.

3. Methodology

The proposed framework integrates deep reinforcement learning (DRL) with robust game-theoretic modeling to optimize financial decision-making in uncertain, multi-agent SCF environments. It systematically defines the state, action, and policy spaces under stochastic and adversarial conditions, enabling the design of adaptive, risk-sensitive, and context-aware financing strategies.

The process begins with the construction of the state space, which reflects the dynamic characteristics of firms and their surrounding macro-financial environment. This is achieved through a state division mechanism based on two key criteria: the enterprise lifecycle stage (e.g., early growth, maturity, or restructuring) and the presence of external uncertainty factors, such as geopolitical risk, inflation, or market volatility. These dimensions jointly define the boundaries of firm behavior under uncertainty and serve as the basis for constructing a stochastic state transition probability matrix, where upper and lower bounds are imposed to reflect volatility constraints.

Following this, the framework defines action space, encompassing the full range of financing and operational choices available to participants in the SCF ecosystem. These actions are embedded in a game-theoretic environment, where heterogeneous decision-makers—including banks, third-party finance platforms, retailers, and manufacturers—interact strategically. The framework models multiple levels of optimization objectives, ranging from individual agent optimization to joint policies involving two, three, or all agents. This allows for the simulation of both bilateral bargaining and system-wide coordination under competing incentives.

To account for environmental instability and adversarial behavior, the model incorporates a robust game theory module, which captures the impact of uncertainty on agent decision-making. This module supports the formulation of robust strategies that anticipate worst-case adversarial scenarios. It also assists in the selection of policy rules that remain valid across a wide range of uncertain conditions. Within this setting, the framework defines the policy space and reward function, where the reward function jointly considers platform profitability, credit risk mitigation, and liquidity stability. Specifically, rewards are structured to balance short-term gains (e.g., interest margins) with long-term resilience (e.g., avoiding borrower default or supply chain disruption).

Once the state transitions, action choices, and reward functions are established, the system enters the training phase, where deep reinforcement learning algorithms are deployed in a simulated environment. Through iterative learning and feedback, the agent progressively refines its policy to maximize expected cumulative rewards. The training process is designed to be adaptive, incorporating environmental feedback to continuously update the strategy space. This stage produces two critical outputs: performance evaluation, which assesses the policy's effectiveness under diverse operational and macroeconomic scenarios, and solution updating, which allows for ongoing refinement and transferability to real-world SCF platforms.

As illustrated in Figure 1, the framework supports a unified decision-support system for intelligent SCF policy design. It integrates credit risk analytics, inter-agent coordination, and robust optimization within a single adaptive learning architecture. The model is capable of handling heterogeneous agent behavior, dynamic state transitions, and system-level liquidity management. Its scalability and interpretability make it suitable for practical deployment in digital finance ecosystems where complexity, uncertainty, and stakeholder heterogeneity are pervasive.

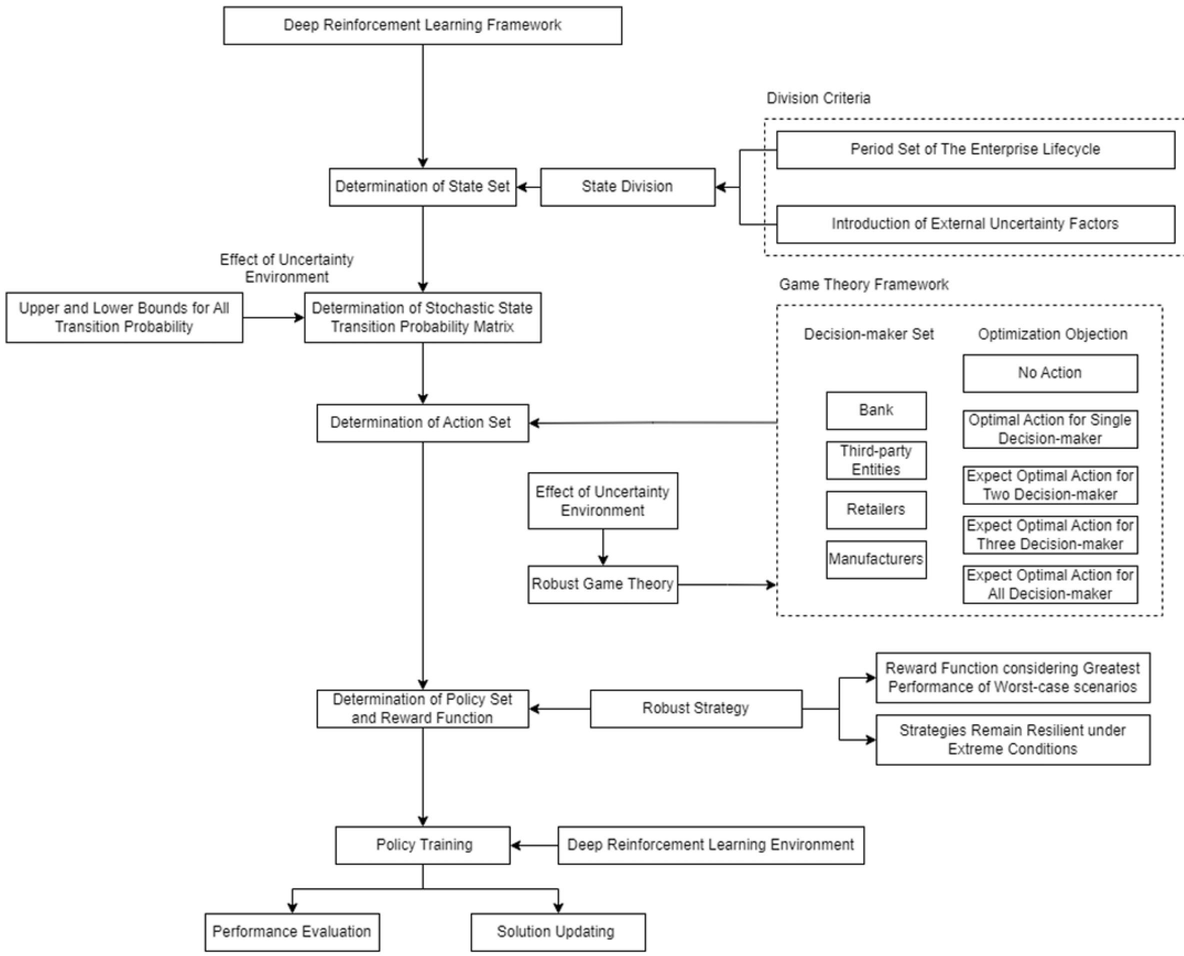

Figure 1. Research framework

4. Numerical Experiments Design

To evaluate the effectiveness of the proposed RL framework in realistic SCF contexts, this study conducts a series of numerical experiments using both synthetic and empirical datasets. The experimental design is intended to assess the framework's robustness, adaptability, and decision quality under a range of macro-financial and operational scenarios. Core performance metrics include credit scoring accuracy, cost efficiency of financing, liquidity optimization, and resilience under adverse conditions.

Two distinct datasets are employed. The first is a synthetic dataset constructed to simulate a heterogeneous population of firms, varying in creditworthiness, liquidity buffers, and roles within the supply chain. This setting enables rigorous testing of policy behavior across controlled yet diverse uncertainty conditions. The second dataset is drawn from empirical sources, namely Eikon and FactSet, and consists of granular financial and transactional data from 100 SMEs operating in the United States and Australia between 2018 and 2022. Variables include quarterly cash flow statements, trade credit terms, invoice payment records, supplier–buyer linkages, and macroeconomic indicators such as inflation, lending rates, and financial stress indices. Data preprocessing involved normalization, temporal alignment, missing value imputation (via forward-fill and k-nearest neighbors), and one-hot encoding of categorical attributes such as industry classification and supplier tier.

The experimental design unfolds in three phases, and each is aligned with a core functionality of the proposed framework. In the first phase, the performance of the credit evaluation module is benchmarked against standard classification models, including logistic regression and random forest classifiers. Evaluation metrics include the area under the ROC curve (AUC), precision-recall curves, and the F1-score for default prediction. The proposed module, comprising gradient boosting and neural network classifiers, achieves superior performance, particularly in identifying early warning signals of liquidity distress among SMEs, which is critical for risk-sensitive SCF strategies.

The second phase evaluates the RL agent's performance in a rolling-horizon, multi-period SCF environment. The RL policy is compared to two benchmark policies:
(i) a static policy using fixed interest rates and fixed credit rules.
(ii) a rule-based policy that adjusts financing terms based on predefined credit score thresholds.

Unlike the benchmarks, the RL agent adapts financing strategies dynamically in response to real-time environmental feedback. Key evaluation metrics in this phase include average financing cost for borrowers, default rate, platform-level profit margin, liquidity ratio, and policy adaptability. The results indicate that the RL policy consistently outperforms both baselines. It achieves an average 15% reduction in financing costs while improving platform profitability and enabling liquidity allocation strategies that better align firm-specific credit conditions with system-wide stability objectives.

The third phase introduces scenario-based stress testing to assess policy resilience under adverse macro-financial shocks. Three stress scenarios are constructed:
(1) a geopolitical shock, characterized by disrupted supply reliability and increased credit demand.
(2) a credit crunch, simulating liquidity tightening and a 25% increase in loan rejection rates.
(3) an inflationary spike, involving sustained increases in interest rates and operational costs.

Each scenario is simulated over 20 Monte Carlo episodes with randomized seed conditions. The RL policy demonstrates robust adaptability, adjusting interest rates, reallocating credit flows, and prioritizing high-credit-quality borrowers with strategic roles in supply chains. By contrast, static and rule-based policies fail to preserve credit continuity in over 40% of runs during the credit crunch scenario, highlighting the limitations of non-adaptive approaches under extreme conditions.

The framework is implemented in TensorFlow 2.0 and trained on an NVIDIA RTX 3090 GPU. Key hyperparameters include a learning rate of 0.0003, a discount factor (γ) of 0.98, a replay buffer of 100,000 transitions, a batch size of 128, and a target network update interval of 10 steps. The RL agent follows an epsilon-greedy exploration strategy, with epsilon annealing from 0.9 to 0.05 over the training horizon. Convergence is typically reached within 5,000 episodes and verified through cumulative reward stabilization and validation consistency across episodes.

In summary, the experimental results provide strong empirical support for the proposed RL-based SCF framework. The model exhibits superior adaptability to volatile financial environments, allocates liquidity based on evolving creditworthiness, and supports credit decision-making aligned with both platform profitability and systemic resilience. It also proves robust under adverse conditions, maintaining credit access and operational stability where static approaches break down. Collectively, these findings underscore the value of RL as a foundation for AI-enhanced SCF platforms, enabling more intelligent, fair, and adaptive financial infrastructures.

5. Results and Analysis

This section presents a detailed comparative evaluation of the experimental outcomes under three policy regimes: the Static Policy (SP), the Rule-Based Policy (RP), and the proposed Reinforcement Learning Policy (RLP). The analysis focuses on four core performance dimensions: financing efficiency, credit risk mitigation, platform profitability, and policy adaptability under stress. In addition, the interpretability of the RL agent is assessed to support regulatory alignment and practical deployment. Figure 2 illustrates the relative performance of each policy across the selected financial metrics.

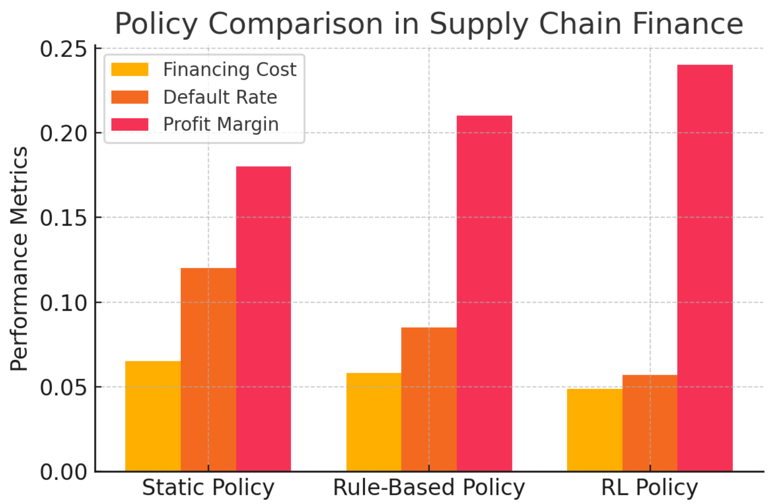

Figure 2. Comparison of policy performance across key financial metrics

5.1 Financing Cost Reduction and Borrower Impact

The RL-based policy demonstrates a significant improvement in reducing the average cost of financing for borrowers. As shown in Figure 2, the RLP achieves a mean interest rate of 4.9%, outperforming the RP (5.8%) and SP (6.5%) baselines. This reduction reflects the RL agent's ability to continuously tailor financing terms based on updated firm-specific credit assessments, real-time liquidity signals, and changing macroeconomic indicators. The agent adjusts interest rates dynamically by learning optimal pricing patterns that balance repayment capacity with market conditions.

This adaptability is particularly beneficial for SMEs, which often face disproportionate financing constraints. By offering risk-adjusted, customized loan terms, the RLP significantly expands access to affordable working capital, even during volatile periods. Furthermore, the policy reduces the frequency of interest rate overestimations for low-risk borrowers and mitigates financial distress that may arise from mismatched loan obligations. This outcome supports both firm-level liquidity management and platform-level loan performance stability.

5.2 Default Rate and Credit Risk Management

The proposed RL framework also demonstrates superior capability in managing credit risk. The RLP yields an average default rate of 5.7%, representing a more than 50% reduction compared to the SP benchmark. This performance stems from the agent's ability to accurately assess default probability in real time and proactively respond to deteriorating financial conditions. The credit module integrates both firm-level indicators (e.g., declining cash flows, late invoice payments) and macro-level shocks (e.g., inflation spikes, currency risk) to generate dynamic risk profiles.

Importantly, the RL policy not only avoids extending credit to high-risk firms but also learns to tighten loan conditions pre-emptively in anticipation of systemic vulnerabilities. For example, under early signs of recessionary pressure, the agent shortens loan maturities and restricts exposure to firms with thin liquidity buffers. This behavior contrasts with RP and SP models, which often continue lending based on static thresholds or historical averages, resulting in higher exposure to latent defaults. Overall, the RLP contributes to stronger portfolio risk containment and proactive risk mitigation.

5.3 Profitability and Platform Performance

From the platform's perspective, the RLP significantly enhances overall financial performance. The policy generates an average profit margin of 24%, compared to 21% under RP and 18% under SP. These gains are attributed to the model's more efficient capital allocation, lower default-related losses, and fine-grained pricing strategies across borrower segments. Rather than applying a one-size-fits-all interest rate structure, the RL agent segments borrowers by credit tier and tailors loan terms to maximize expected return per unit of risk.

The model's pricing logic reflects real-time shifts in borrower behavior and market liquidity, allowing it to prioritize lending to firms that contribute most effectively to ecosystem stability and revenue. Moreover, the reduction in loss provisioning (due to fewer defaults) further improves profitability. The rule-based policy, in contrast, fails to adapt pricing logic dynamically, while the static policy lacks any feedback-driven optimization mechanism, leading to inefficient interest spreads and suboptimal return-risk trade-offs.

5.4 Policy Adaptability Under Stress

Stress testing results confirm the RLP's robust adaptability in volatile financial conditions. In the credit crunch scenario, characterized by tightening liquidity and increased loan rejection pressure, the RL agent successfully prioritizes funding for strategically important suppliers, typically those with shorter repayment cycles, strong cash conversion efficiency, and upstream integration value. This ensures continuity of core supply chains while reducing exposure to risk-prone firms.

In the inflationary shock scenario, the RL policy adjusts interest rates incrementally in response to rising capital costs, avoiding sudden price hikes that would deter borrowing or cause demand collapse. This contrasts sharply with RP and SP, which often either overreact (leading to unnecessary credit rationing) or underreact (resulting in erosion of platform capital base). Furthermore, the RL agent demonstrates multi-period learning behavior, refining its responses over time as macro conditions evolve. These adaptive dynamics highlight the policy's capacity to preserve liquidity flow and operational resilience in high-volatility environments where static models systematically fail.

5.5 Policy Interpretability and Strategic Insights

Beyond performance metrics, interpretability is crucial for stakeholder acceptance and regulatory alignment. The RLP integrates Shapley value-based feature attribution to explain key decision drivers behind loan approval and pricing recommendations. Analysis shows that firm-specific credit scores, liquidity buffer ratios, and macroeconomic variables such as volatility indices and interest rate trends are consistently among the most influential factors.

These findings not only enhance model transparency but also provide valuable insights for product innovation and risk pricing. For instance, in high-risk environments, the model consistently favors short-term financing structures with capped rates, reflecting a learned preference for controlling duration and price volatility. Such behavior can inform the design of new lending instruments for financially vulnerable firms. The interpretability module also allows compliance teams and platform managers to audit, validate, and adjust policy decisions, improving accountability and governance in AI-driven SCF systems.

6. Conclusion

This study contributes to SCF literature by introducing an RL-based decision framework that integrates dynamic credit assessment, multi-agent coordination, and adaptive policy optimization under uncertainty. The proposed model formulates SCF as a stochastic Markov Decision Process, enhanced by game-theoretic reasoning, to support intelligent, context-aware financial decision-making that evolves with both firm-level and macroeconomic conditions. Empirical validation conducted using both synthetic simulations and real-world data from multinational corporations demonstrates that the RL policy consistently outperforms static and rule-based benchmarks. Notable improvements are observed in financing cost reduction, platform-level profitability, and systemic resilience, particularly in stress scenarios involving liquidity constraints and geopolitical shocks.

Theoretically, this work advances the field of AI-enabled finance by embedding learning-based decision mechanisms within complex, multi-agent SCF systems. Practically, it offers actionable insights into improving credit accessibility, liquidity allocation, and risk management, especially for financially constrained SMEs. The integration of explainability tools—such as Shapley-based attribution—also enhances transparency, stakeholder trust, and regulatory compliance.

Nevertheless, certain limitations remain. These include assumptions regarding agent behavior, potential data sparsity, and challenges in generalizing across diverse institutional contexts. Future research should explore MARL architectures, IoT-integrated environments, smart contract automation, and fairness-aware optimization techniques to further improve the robustness, scalability, and inclusiveness of digital SCF systems.

References

Cowton, C. J., and San-Jose, L., Trade credit optimization under uncertainty, 2016.

Deng, Y., Bao, F., Kong, Y., Ren, Z., and Dai, Q., Deep direct reinforcement learning for financial signal representation and trading, IEEE Transactions on Neural Networks and Learning Systems, 2016.

Dong, Y., et al., Blockchain applications in supply chain finance visibility, 2023.

Fischer, T., Reinforcement learning in financial markets—A survey, Advances in Financial Machine Learning, 2018.

Gelsomino, L. M., et al., Supply chain finance literature review, International Journal of Physical Distribution and Logistics Management, 2016.

Jiang, Z., Xu, D., and Liang, J., A deep reinforcement learning framework for the financial portfolio management problem, arXiv preprint arXiv:1706.10059, 2017.

Kouvelis, P., and Zhao, W., External financing and supply chain finance policy design, 2018.

Li, J., Deng, Y., and Zhang, J., Reinforcement learning applications in finance: A survey, ACM Computing Surveys, 2019.

Luo, Y., and Witt, M. A., Springboard MNEs under de-globalization: The case of China's Belt and Road Initiative, Journal of International Business Studies, vol. 53, no. 4, pp. 767–780, 2021.

Pei, Q., Chan, H. K., Zhang, T., and Li, Y., Benefits of the implementation of supply chain finance, Annals of Operations Research, vol. 331, no. 1, pp. 251–283, 2022.

Rajagopal, M., et al., Artificial intelligence in supply chain finance, ICONSTEM, 2023.

Ronchini, A., Guida, M., Moretto, A., and Caniato, F., The role of artificial intelligence in the supply chain finance innovation process, Operations Management Research, 2024.

Sutton, R. S., and Barto, A. G., Reinforcement learning: An introduction, 2nd ed., MIT Press, 2018.

Tanrisever, F., et al., Working capital optimisation in reverse factoring, 2015.

Wang, T., Liu, X., and Zhu, Y., Q-learning in supply chain finance, Expert Systems with Applications, vol. 222, art. 119231, 2023.

Biographies

Yehan Dou received her Bachelor of Engineering degree in Logistics Engineering from Southwest Jiaotong University. She completed her Master's degree in Logistics Engineering and Management at Tsinghua University, where she developed a strong academic foundation in supply chain systems and operational decision-making. She is currently a PhD candidate at the Macquarie Business School, Macquarie University, Australia. Her research interests lie in supply chain management, supply chain optimization, and operations research, with a particular focus on data-driven modeling and intelligent decision support in complex logistics and financial networks. Her doctoral work aims to enhance the resilience and efficiency of supply chains through advanced optimization and AI-enabled methods.

Dr. Peter Shi is a Senior Lecturer in the Department of Actuarial Studies and Business Analytics at Macquarie University. He currently serves as Deputy Course Director and Head of the Supply Chain Management Specialization within the Master of Commerce program. His expertise lies in operations and supply chain management, with a strong focus on stochastic optimization, digital transformation, logistics systems, and cross-border e-commerce. He has published extensively in leading peer-reviewed journals, including Annals of Operations Research, Production Planning and Control, and Transportation Research Part E. His recent research explores the applications of machine learning, blockchain, and artificial intelligence in sustainable supply chains, logistics innovation, and responsible sourcing. He has led and contributed to interdisciplinary research projects on digital banking transformation and market-oriented elderly care services. He supervises PhD and Master's students working on topics such as supply chain finance, technology adoption in emerging economies, and strategic logistics decision-making. Actively engaged in the academic and professional community, Dr. Shi serves as Chairman of the Chartered Institute of Logistics and Transport (CILT) – New South Wales Chapter. He is Associate Editor of Business Strategy and the Environment and Area Editor for Logistics and Supply Chain Management at Computers and Industrial Engineering. In addition, he serves on the editorial review boards of several high-ranking journals, including International Journal of Logistics Management, International Journal of Operations and Production Management, International Journal of Physical Distribution and Logistics Management, Journal of Purchasing and Supply Management, and Transportation Research Part E. He is a Chartered Fellow of the Chartered Institute of Logistics and Transport and has received multiple recognitions for his research and editorial contributions, including awards for his work on blockchain applications in logistics and systematic reviews in sustainable supply chain strategy. His international collaborations span more than 30 countries and regions, including China, India, the United States, Canada, France, Singapore, and New Zealand.

Ruoyu Li is a PhD candidate in the School of Engineering and Information Sciences at the University of Wollongong. He received his Bachelor's degree in Logistics Engineering from Southwest Jiaotong University and completed his Master's degree in Logistics Engineering at Nanjing University. His current research focuses on the intersection of machine learning and operations research, with particular interests in intelligent decision-making, algorithmic optimization, and data-driven approaches to logistics and supply chain systems. His doctoral work explores the development of hybrid machine learning and mathematical programming techniques to solve large-scale optimization problems under uncertainty. Ruoyu Li aims to contribute to the advancement of intelligent, robust, and efficient decision-support systems for complex engineering and logistics applications.