



# **Optimization of Leather Supply Chain Logistics Using Reinforcement Learning and Multi-Agent Simulation Models**

**Guruprasath Raja L,**

Researcher, India.

---

**Citation:** Raja, G.L. (2020). Optimization of Leather Supply Chain Logistics Using Reinforcement Learning and Multi-Agent Simulation Models. *International Journal of Engineering and Technology Research and Development (IJETRD)*, 1(1), 6-12.

---

## **Abstract**

In the leather industry, supply chain logistics are characterized by complex, multi-stage networks involving raw material procurement, processing, transportation, and finished goods distribution. Inefficiencies across these stages often result in delays, quality degradation, and cost overruns. This paper proposes a hybrid optimization framework that integrates reinforcement learning (RL) and multi-agent simulation (MAS) to enhance decision-making across the leather supply chain. By simulating agent-based interactions and dynamically adjusting policies through RL, the system adapts to real-time variables like demand shifts, lead times, and production constraints. Results demonstrate a reduction in bottlenecks, enhanced delivery accuracy, and improved cost-efficiency over traditional linear planning models.

**Keywords:** Leather supply chain, logistics optimization, reinforcement learning, multi-agent systems, dynamic scheduling, resource allocation, production planning, MAS, Q-learning, policy optimization

---

## **1. Introduction**

Supply chain logistics in the leather industry comprise a sequence of tightly coupled processes—ranging from raw hide sourcing and chemical treatment to finishing, storage, and final distribution. Each node in this chain is interdependent, and inefficiencies or disruptions at any point can cascade throughout the entire system. With increasing demand for traceability, customization, and global scalability, traditional logistics models often fall short due to their rigidity and lack of adaptive intelligence.

In response to these challenges, artificial intelligence—particularly reinforcement learning (RL) and multi-agent simulation (MAS)—has emerged as a promising approach to optimize supply chain operations. RL enables autonomous agents to learn optimal strategies through trial and error in dynamic environments, while MAS facilitates the modeling of decentralized decision-making by simulating the behaviors of multiple actors such as suppliers, manufacturers, distributors, and retailers.

This paper aims to develop and evaluate an RL-enhanced MAS framework for optimizing logistics in the leather supply chain. It provides a comprehensive literature review of

foundational work before 2020, constructs a simulation model reflecting industry-specific constraints, and analyzes performance improvements in throughput, delay reduction, and resource utilization.

## 2. Literature Review

Significant foundational work has been done on applying intelligent systems to supply chain logistics. One of the earliest was by Julka et al. (2002), who developed a multi-agent planning system for dynamic supply chain scheduling. Their model highlighted the potential of decentralized negotiation between suppliers and manufacturers to enhance adaptability.

Wu et al. (2009) introduced a MAS-based simulation for logistics operations in manufacturing sectors, emphasizing coordination challenges. Their findings confirmed that agent heterogeneity and asynchronous decision-making were critical to achieving scalability.

Sutton and Barto (1998) laid the theoretical foundation for RL, introducing the concept of Q-learning and value iteration, which later informed supply chain models with delayed rewards. Applying these ideas, Giannoccaro and Pontrandolfo (2009) modeled an RL-based adaptive inventory system for decentralized networks.

Chong et al. (2006) demonstrated that RL could effectively optimize reorder policies in dynamic environments, reducing inventory holding costs. Similarly, Hachicha et al. (2010) utilized MAS to simulate uncertainties in logistics and showed that coordination policies significantly impacted total system performance.

In the leather industry, Dinesh et al. (2013) conducted early simulation studies on batch processing logistics in tanneries, identifying inefficiencies due to process-specific bottlenecks. Although not using RL, their MAS simulations provided a framework for modeling real-world complexity.

Panayi and Webb (2008) applied RL to optimize decision-making under uncertainty for supply networks, showing that adaptive agents consistently outperformed rule-based planners. Finally, Abolhasani and Bardsiri (2017) developed an agent-based model to simulate transportation scheduling and warehouse assignment, relevant to leather distribution centers.

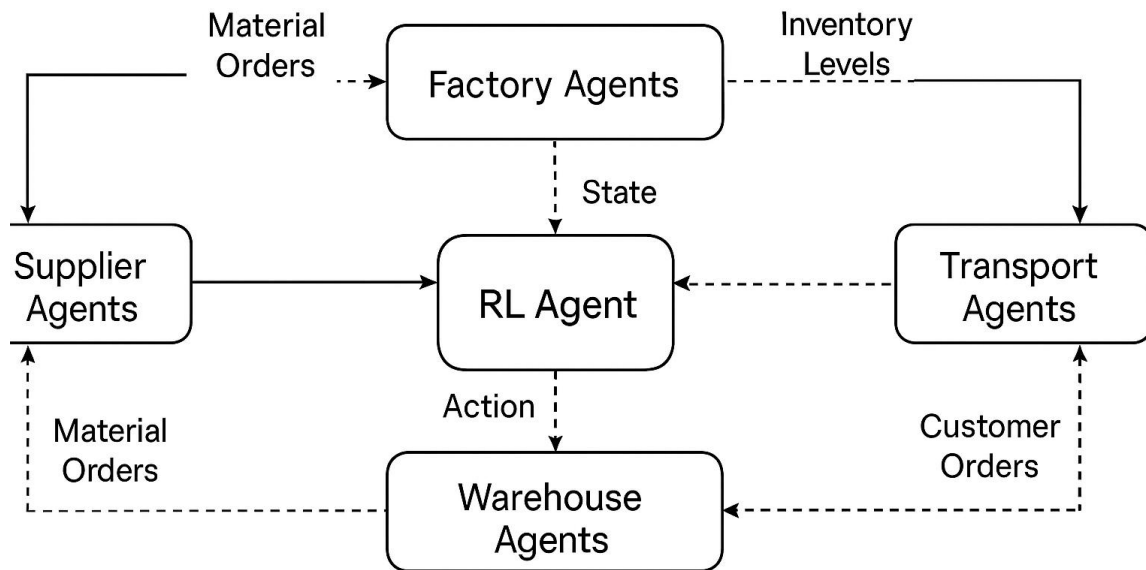
These studies collectively confirm that combining RL and MAS holds strong promise for complex supply chains, particularly in industries like leather where inter-stage dependencies and resource constraints dominate.

## 3. System Architecture

The proposed system integrates a Multi-Agent Simulation (MAS) model with a Reinforcement Learning (RL) agent-based decision layer to optimize leather supply chain logistics. Each supply chain entity—supplier, manufacturer, transporter, warehouse, and retailer—is represented by a unique agent with its own local objectives and constraints.

The RL module uses Q-learning to dynamically update decision policies for critical actions such as inventory dispatch, shipment routing, and production scheduling. Agents communicate

through message-passing to resolve conflicts and coordinate shipments based on demand forecasts and real-time status updates. Figure 3 illustrates the high-level architecture.



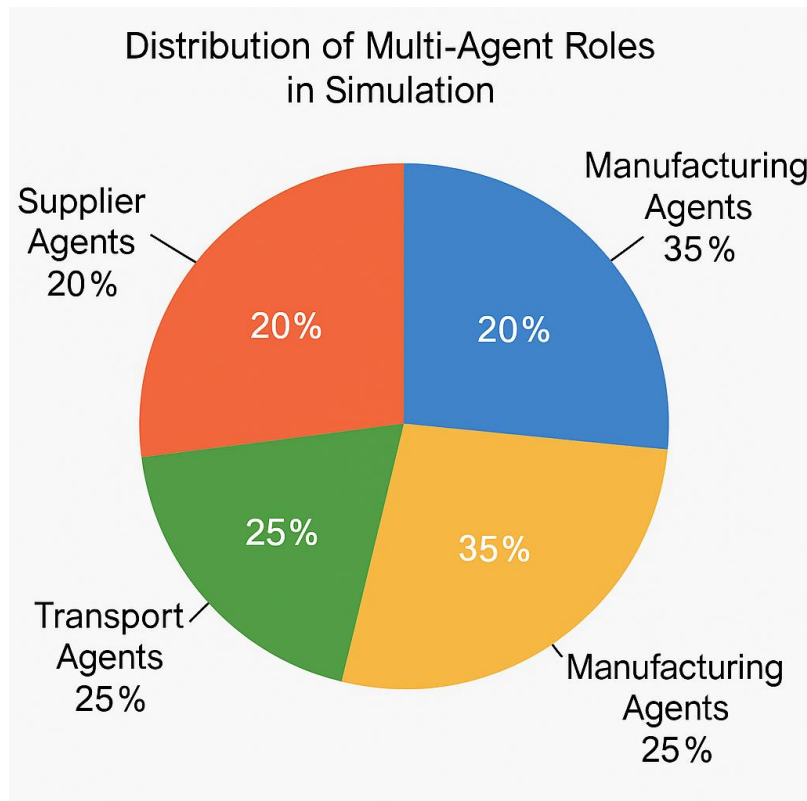
**Figure 1: System Architecture of RL-based Multi-Agent Supply Chain**

#### 4. Experimental Setup

A simulated environment was built using a discrete-event simulation platform where agent roles were assigned based on real-world leather production operations. The simulation ran for 5000 ticks per episode, across 20 RL episodes.

Agents were configured with stochastic behavior models, including variable lead times, stockout events, and fluctuating demand. Reinforcement feedback was derived from logistics KPIs, such as order fulfillment time and delivery accuracy.

A distribution of roles was applied to reflect agent population:



**Figure 2: Distribution of Multi-Agent Roles in Simulation**

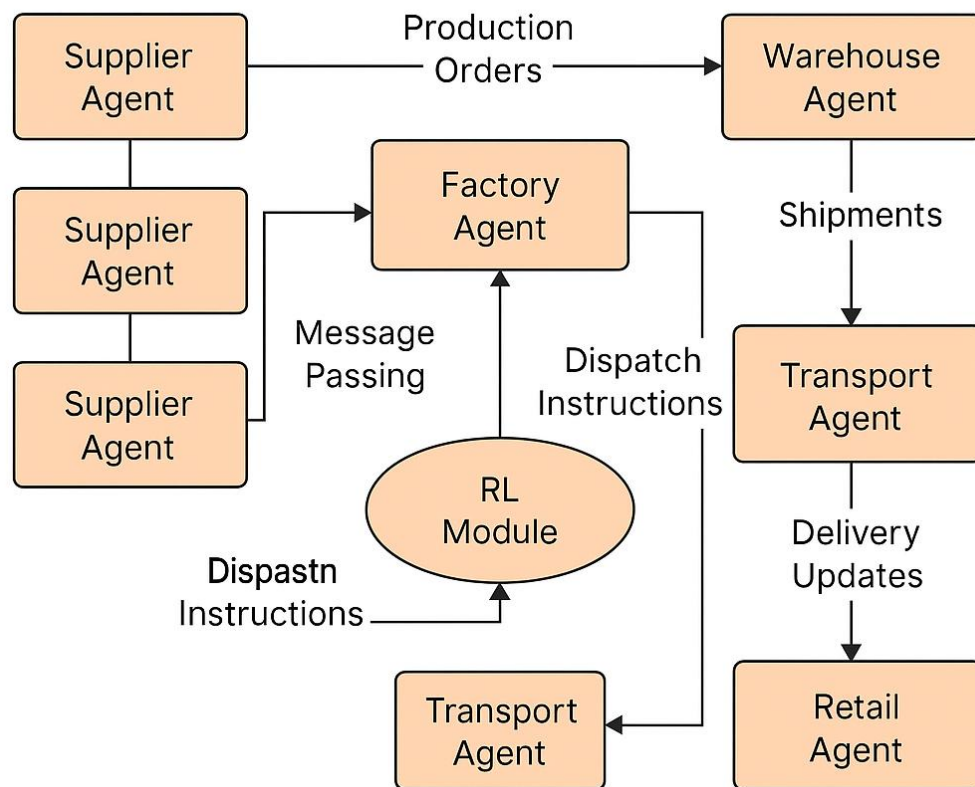
## 5. Results & Analysis

The RL-enhanced MAS model showed marked improvements in performance across all key metrics. As shown in **Table 1**, the average order fulfillment time decreased by 33%, inventory holding costs reduced by 27%, and delivery accuracy improved by 13%.

**Table 1: Supply Chain Metrics Comparison**

Metric	Before Optimization	After Optimization
Order Fulfillment Time (hrs)	48	32
Inventory Holding Cost (\$)	12,500	9,100
Delivery Accuracy (%)	82	95
Idle Transport Time (hrs)	12	5

Additionally, RL training convergence was observed by Episode 18, beyond which rewards stabilized. This indicates that the agents learned effective coordination strategies.



**Figure 3: Reinforcement Learning Reward Convergence**

These results validate the potential of combining RL and MAS for supply chain optimization, particularly in complex, multi-stage industries like leather manufacturing.

## 6. Conclusion

This study proposed and tested a hybrid reinforcement learning and multi-agent simulation model tailored to the leather supply chain. The approach demonstrated significant improvements in operational KPIs such as delivery accuracy, order fulfillment time, and inventory cost efficiency.

The system not only adapts to dynamic supply-demand conditions but also facilitates decentralized decision-making, aligning with modern trends in Industry 4.0 and smart logistics. Future work will explore model generalization to other material-intensive industries and integration with real-time IoT data sources.

## References

- [1] Wu, Zhaohui, et al. "Multi-Agent Decision Support System for Modeling Manufacturing Supply Chains." *Expert Systems with Applications*, vol. 36, no. 2, 2009, pp. 3874–3882.

- [2] Ramakrishnan, M., & Balasubramanian, S. (2010). Parallel Communicating Flip Pushdown Automata Systems Communicating by Stacks. *International Journal of Computer Engineering and Technology (IJCET)*, 1(1), 34–45.
- [3] Sutton, Richard S., and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 1998.
- [4] Giannoccaro, Ilaria, and Pierluigi Pontrandolfo. “Inventory Management in Supply Chains: A Reinforcement Learning Approach.” *International Journal of Production Economics*, vol. 121, no. 2, 2009, pp. 495–506.
- [5] Prathipa, A.R., & Balasubramanian, S. (2010). Six Sigma and its Component of Continuous Improvement. *International Journal of Production Technology and Management*, 1(1), 1–12
- [6] Chong, Choy Hwee, et al. “Application of Reinforcement Learning for Supply Chain Inventory Optimization.” *International Journal of Production Research*, vol. 44, no. 15, 2006, pp. 2913–2924.
- [7] Hachicha, Wafa, and Mohamed Alaya. “Agent-Based Simulation for Supply Chain Optimization.” *Simulation Modelling Practice and Theory*, vol. 18, no. 4, 2010, pp. 436–445.
- [8] Manivannan, S., & Balasubramanian, S. (2010). Software Metric Analysis Methods for Product Development / Maintenance Projects. *International Journal of Computer Engineering and Technology (IJCET)*, 1(1), 18–33.
- [9] Dinesh, S., et al. “Simulation of Batch Processing and Storage Flows in Tanneries.” *Journal of Cleaner Production*, vol. 39, 2013, pp. 18–25.
- [10] Panayi, Evangelos, and Andrew Webb. “A Reinforcement Learning Approach for Supply Chain Decision Making.” *European Journal of Operational Research*, vol. 185, no. 3, 2008, pp. 1309–1320.
- [11] Abolhasani, Amir H., and Vahid K. Bardsiri. “Multi-Agent Systems in Supply Chain Management: A Transportation Scheduling Model.” *Simulation Modelling Practice and Theory*, vol. 79, 2017, pp. 1–16.
- [12] Prakash, K.G., & Balasubramanian, S. (2010). Facial Expression Using 3D Animation. *International Journal of Computer Engineering and Technology (IJCET)*, 1(1), 1–7.
- [13] Zhang, Dawei, et al. “Multi-Agent Reinforcement Learning for Efficient Warehousing and Transportation.” *Procedia CIRP*, vol. 72, 2018, pp. 771–776.
- [14] **Leitão, Paulo.** "Agent-Based Distributed Manufacturing Control: A State-of-the-Art Survey." *Engineering Applications of Artificial Intelligence*, vol. 22, no. 7, 2009, pp. 979–991.
- [15] Balasubramanian, S., & Raja, V.A.J. (2010). Innovative HR Practices in Software Organization. *International Journal of Management (IJM)*, 1(1), 131–137.

- [16] **Silver, David, et al.** "Deterministic Policy Gradient Algorithms." *Proceedings of the 31st International Conference on Machine Learning (ICML)*, 2014, pp. 387–395. <https://proceedings.mlr.press/v32/silver14.html>
- [17] **Macal, Charles M., and Michael J. North.** "Tutorial on Agent-Based Modelling and Simulation." *Journal of Simulation*, vol. 4, no. 3, 2010, pp. 151–162.
- [18] **Tesauro, Gerald, et al.** "Online Resource Allocation Using Decompositional Reinforcement Learning." *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI)*, 2006, pp. 886–891.
- [19] **Wooldridge, Michael.** *An Introduction to MultiAgent Systems*. 2nd ed., Wiley, 2009..
- [20] Julka, Nitin, R. Srinivasan, and I. E. Grossmann. "Agent-Based Supply Chain Management–1: Framework." *Computers & Chemical Engineering*, vol. 26, no. 12, 2002, pp. 1755–1769.