Annals of the Academy of Romanian Scientists
Series on Engineering Sciences
ISSN 2066-8570 Volume **17**, Number **2**/**2025**

MONTE CARLO MARKOV CHAIN (MCMC) STOCHASTIC MODELING OF SUPPLY CHAIN

Marcel ILIE[1] and Augustin SEMENESCU[2]

Rezumat. *Gestionarea eficientă a stocurilor în lanțurile de aprovizionare multi-echelon se confruntă cu provocări semnificative din cauza cererii aleatorii și a incertitudinii timpilor de livrare, care amplifică variabilitatea și cresc costurile operaționale. În acest studiu, propunem un cadru bazat pe lanțuri Markov pentru modelarea, analizarea și optimizarea sistemelor de inventar multi-echelon în condiții de timpi de livrare stocastici. Cadrul conceptualizează nivelurile stocurilor și stările timpilor de livrare ca un sistem de tranziții probabilistice, permițând calculul distribuțiilor în stare staționară, al probabilităților de epuizare a stocului și al timpilor de recuperare tranzitorie. Rezultatele analitice sunt validate prin simulări Monte Carlo, evidențiind o concordanță ridicată între distribuțiile teoretice și cele empirice. Experimentele numerice arată impactul variabilității timpilor de livrare asupra performanței stocurilor, relevând creșteri neliniare ale probabilității de lipsă de stoc și ale costului total al sistemului pe măsură ce varianța timpului de livrare crește. Analizele multi-echelon ilustrează apariția efectului „bullwhip" și subliniază eficiența partajării informațiilor în reducerea propagării variabilității între echeloane. Comparativ cu politicile bazate pe învățarea prin consolidare profundă (DRL), se constată că, deși DRL poate atinge costuri totale ușor mai reduse, abordarea bazată pe Markov oferă avantaje semnificative în ceea ce privește interpretabilitatea, robustețea și eficiența computațională. Studiul contribuie teoretic prin unificarea dinamicii stocastice multi-echelon și a analizei tranzitorii într-un cadru Markov tratabil. Din perspectivă managerială, oferă recomandări aplicabile pentru reducerea varianței timpilor de livrare, creșterea vizibilității între echeloane și proiectarea de politici hibride analitice–de învățare. Acest cadru constituie astfel o bază solidă pentru controlul stocurilor rezilient și eficient din punct de vedere al costurilor în rețele complexe și incerte de aprovizionare.*

Abstract. *Effective inventory management in multi-echelon supply chains is challenged by stochastic demand and uncertain lead times, which amplify variability and increase operational costs. This study presents a Markov chain framework for modeling, analyzing, and optimizing multi-echelon inventory systems under stochastic lead-time conditions. The framework represents inventory levels and lead-time states as a probabilistic transition system, enabling computation of steady-state distributions, stockout probabilities, and transient recovery times. Analytical results are validated through Monte Carlo simulations, demonstrating high fidelity between theoretical and empirical distributions. Numerical experiments quantify the impact of lead-time variability on inventory performance, revealing nonlinear increases in stockout*

[1]Marcel ILIE, Associate. Prof. Ph.D. Georgia Southern University, 1332 Southern Dr. Statesboro GA 30458, USA, *Corresponding author:, milie@georgiasouthern.edu

[2] Augustin SEMENESCU, Prof. National Science and Technology University Politehnica Bucharest, Spl.Independentei 313, Bucharest, Romania, augustin.semenescu@upb.ro

probability and total system cost as lead-time variance grows. Multi-echelon analyses demonstrate the emergence of the bullwhip effect and highlight the effectiveness of information sharing in mitigating variability propagation across echelons. Comparative benchmarking against deep reinforcement learning (DRL) policies shows that while DRL achieves marginally lower total costs, the Markov-based approach provides superior interpretability, robustness, and computational efficiency. The study offers theoretical contributions by unifying stochastic multi-echelon dynamics and transient analysis within a tractable Markov framework. Managerially, it provides actionable insights on lead-time variance reduction, cross-echelon visibility, and hybrid analytical–learning policy design. The framework establishes a foundation for resilient and cost-effective inventory control in complex, uncertain supply-chain networks.

Keywords: multi-echelon inventory, stochastic lead time, Markov chain, supply chain management, bullwhip effect, reinforcement learning

DOI 10.56082/annalsarscieng.2025.2.94

1. Introduction

Modern supply chains are increasingly global, multi-tiered and tightly coupled: components and finished goods routinely traverse several echelons (suppliers → manufacturers → distributors → retailers) before reaching the final customer. This structural complexity interacts with volatile demand patterns, transportation variability, supplier disruptions and geopolitical events to produce pronounced operational risk. One of the most consequential forms of operational uncertainty is lead-time variability-the stochastic delay between placing an order and receiving replenishment-which amplifies inventory risk (higher safety stocks, more frequent stockouts) and complicates coordination across echelons. Recent surveys and reviews identify lead-time uncertainty as a central research and practice challenge that continues to motivate new analytical and computational methods.

Multi-echelon networks magnify the problem. Inventory decisions at one echelon affect upstream replenishment timing and downstream service: for example, an upstream stockout can lengthen the effective lead time faced by downstream facilities, while abnormal lead times at a supplier change the optimal safety-stock allocation throughout the network. Classic single-stage or decoupled approaches therefore often give misleading performance predictions when applied to multi-echelon systems. Literature reviews and modelling studies show that multi-echelon inventory design-including safety-stock placement, reorder policy coordination and lead-time modelling-is essential to capture real network behavior.

Against this backdrop, stochastic dynamic models that explicitly represent state evolution (inventory on-hand, outstanding orders, shipments in transit, and lead-time phases) have shown promise. Markov chain and Markov decision process (MDP) formulations are particularly well-suited: they represent the system as a discrete (or discretized) state space with probabilistic transitions, enabling exact or approximate analysis of steady-state distributions, transient behaviors, and policy optimization under uncertainty. These models provide a principled route to quantify trade-offs (holding vs. shortage costs), evaluate stocking and reorder thresholds, and incorporate state-dependent lead-time processes. The growing use of MDPs and learning-based methods for multi-echelon problems (including deep reinforcement learning) illustrates both the feasibility and current interest in state-based stochastic frameworks. Nevertheless, three persistent gaps remain in the literature: (1) relatively few tractable Markov-chain models capture stochastic lead-time processes that depend on upstream states and network interactions, (2) many studies focus on small or stylized networks (one or two echelons) or assume fixed lead times, limiting applicability to modern, interconnected supply networks, and (3) scalable computation of steady-state metrics and policy optimization in realistic multi-echelon state spaces remains an open challenge. Recent reviews emphasize the need for modelling frameworks that combine fidelity (capturing lead-time dynamics and echelon dependencies) with computationally tractable solution methods.

This paper addresses those gaps by developing a Markov-chain-based modelling and analysis framework for multi-echelon inventory systems under lead-time uncertainty.

Our main contributions are:

- Modeling: a compact but expressive state-space representation that tracks on-hand inventory, outstanding orders, and lead-time phases across multiple echelons; lead times are modelled as stochastic processes whose transition probabilities can depend on upstream states (e.g., supplier availability).
- Analysis: methods to compute stationary distributions (when they exist) and closed-form or numerically tractable expressions for performance metrics (average inventory, stockout probability, fill rate, expected cost), including dimension-reduction approaches to keep computation practical.
- Policy insights: comparative experiments showing how lead-time variability, echelon count and policy parameters (reorder levels, coordination rules) affect cost-service trade-offs and where investments -e.g., lead-time reduction vs. inventory rebalancing-produce the largest benefits.

- Extensions & implementation: demonstration of how the Markov framework can integrate simple learning/MDP updates or be combined with approximate dynamic programming (or RL) for larger networks.

By combining explicit lead-time stochasticity with multi-echelon state coupling and by emphasizing computationally feasible analysis, the framework offers both theoretical insight and practical guidance for inventory policy design under realistic uncertainty.

The remainder of this paper is organized as follows. Section 2 surveys related work on multi-echelon inventory systems, stochastic lead times, and Markov/MDP approaches. Section 3 presents the model, assumptions and mathematical formulation. Section 4 describes numerical experiments, computational methods and sensitivity analyses. Section 5 discusses managerial implications, limitations and potential extensions. Section 6 concludes.

2. Literature review

2.1 Classical Multi-Echelon Inventory Theory

The study of multi-echelon inventory systems originated with the foundational work of Clark and Scarf, 1960 [5], who demonstrated that optimal policies in serial multi-stage systems could be decomposed into equivalent single-stage problems through the use of echelon stock concepts. Their analysis established that under convex cost structures and serial dependencies, base-stock policies are optimal. These results laid the theoretical foundation for subsequent multi-echelon research. Subsequent studies extended the Clark–Scarf framework to accommodate more complex supply-chain topologies, stochastic demand processes, and batch ordering constraints [3, 4,]. Such extensions retained key structural insights-such as monotonicity and convexity-that allow efficient computation of near-optimal inventory policies. These classical contributions remain influential, as they continue to guide both analytical and heuristic approaches to inventory management in large-scale supply networks.

2.2 Lead-Time Uncertainty in Supply Chains

Uncertainty in lead time-the delay between ordering and receiving inventory-is widely recognized as a major driver of supply-chain risk and inefficiency. Variability in lead times can stem from production delays, transportation congestion, customs inspections, or supplier reliability issues [15]. Empirical evidence indicates that higher lead-time variance increases required safety stock levels and decreases service performance [8]. Classical inventory models often assume fixed lead times, which limits their realism for globalized, multi-echelon systems where delays fluctuate dynamically. To capture these effects, researchers have developed several stochastic frameworks. One common approach models lead times as independent random variables, while others employ Markov-modulated lead-time processes to capture temporal correlations or regime shifts [2]. More advanced models incorporate state-dependent lead times, where replenishment delays depend on upstream inventory or capacity states [13, 14]. These formulations provide richer representations of real-world uncertainty at the cost of increased analytical complexity.

2.3 Markov Chain and Markov Decision Process Approaches

Markov chain and Markov decision process (MDP) models provide a powerful mathematical framework for representing the stochastic evolution of inventory systems. In these models, the system state includes on-hand inventory, outstanding orders, and lead-time phases; transitions between states occur with known probabilities [16]. Such formulations enable the derivation of long-run metrics such as steady-state probabilities, expected costs, and stockout risks. For single-stage systems, optimal ordering policies often exhibit simple threshold structures-commonly base-stock or (s, S) rules [11]. Extending these insights to multi-echelon systems, however, introduces significant computational challenges because the state space grows exponentially with the number of echelons and stochastic lead-time phases. To address this, a single-unit decomposition approach that breaks a large multi-echelon problem into smaller, tractable sub-problems with Markov-modulated demand and lead-time dynamics was proposed [10]. Their approach demonstrated that near-optimal policies can be derived efficiently while preserving key stochastic dependencies.

2.4 Approximation and Decomposition Methods

Given the computational intractability of solving exact MDPs for large networks, researchers have developed various approximations and decomposition methods. A multi-echelon inventory positioning model that optimizes safety-stock placement using approximate lead-time demand distributions was also proposed [6]. Other methods use base-stock heuristics or mean-field approximations to simplify interactions between echelons while maintaining acceptable accuracy [1].

Simulation-based optimization techniques have also gained attention for estimating steady-state performance when analytical solutions are unavailable. Such methods can incorporate realistic features-nonstationary demand, correlated lead times, and partial observability-but may require significant computational effort to converge [14]

2.5 Emerging Learning-Based and Computational Approaches

Recent years have witnessed growing interest in using reinforcement learning (RL) and deep reinforcement learning (DRL) to optimize multi-echelon inventory control. RL frameworks learn policies directly from system interactions, avoiding the need for explicit analytical transition models [9]. It was demonstrated that DRL can effectively coordinate ordering policies across multiple echelons, outperforming classical heuristics under nonstationary and uncertain lead-time conditions [6]. Other studies have explored multi-agent RL formulations, in which each echelon learns locally while coordinating through shared reward structures [12]. These approaches are promising for high-dimensional problems but face challenges related to convergence stability, interpretability, and performance guarantees. As such, hybrid frameworks that integrate Markov-chain models with learning-based control are emerging as a compelling direction for future research [2, 15].

2.6 Research Gaps and Motivation

Although the literature on stochastic inventory management is extensive, several important gaps remain. First, few existing models jointly capture multi-echelon dependencies and state-dependent lead-time uncertainty within a unified Markovian framework. Second, computational methods for such models remain limited in scalability, restricting their applicability to realistic industrial systems. Third, there is a lack of comparative analysis between analytical Markov methods and data-driven learning approaches for multi-echelon networks. This study addresses these gaps by developing a Markov-chain-based framework for multi-echelon inventory management that explicitly models stochastic lead times and provides tractable steady-state and performance analyses. The approach bridges classical stochastic inventory theory with modern computational tools, providing both theoretical and managerial contributions.

3. Methodology

3.1 Overview of the Stochastic Modeling

In supply chain systems, uncertainty in demand, lead times, and supplier reliability can lead to stockouts, overstock, or disrupted flows. A Markov chain model offers a mathematically tractable way to model the probabilistic transitions between discrete supply chain states over time. In this methodology, we describe how to define states, estimate transition probabilities, validate the model, and conduct scenario/sensitivity analysis in a supply chain context.

3. 2 System Definition and State Space Construction

3. 2. 1 Supply Chain Variables and Discretization

To apply a Markov chain model, we first select key observable variables that capture system behavior-for example:

- Inventory level at a key node (low/medium/high)
- Demand regime (e.g. low, normal, high)
- Supplier performance (on-time, delayed)
- Lead time variability regime (low variance, high variance)

These continuous or semi-continuous observations are discretized into finitely many states by domain knowledge or clustering (e.g. quantiles or k-means). Let the resulting state space be

$$S = \{s_1, s_2, \dots, s_n\} \tag{1}$$

each representing a distinct operational regime

(e.g. $s_1 =$ "$low\ inventory, high\ demand, delaed\ supplier$)

3.2 State Transition Assumption

We assume a discrete-time Markov chain, with time index (t = 0, 1, 2…). The chain has the Markov property:

$$P\left(X_{t+1} = s_j \middle| X_t = s_i, X_{t-1}, \dots\right) = P\left(X_{t+1} = s_j \middle| X_t = s_i\right) \tag{2}$$

Thus, the future state depends only on the current state. The transition structure is represented by a transition probability matrix:

$$P = [p_{ij}] \tag{3}$$

$$p_{ij} = P(X_{t+1} = s_j | X_t = s_i) \text{ , } i, j = 1, \dots, n, \tag{4}$$

with each row summing to 1:

$$\sum_{j=1}^{n} p_{ij} = 1 \tag{5}$$

4. Estimation of Transition Probabilities

4.1Empirical Frequency Estimation

Using historical data over discrete periods, we count observed transitions:

- Let n_{ij} be the number of observed transitions from state s_i to s_j
- Then the empirical estimator is

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{j=1}^{n} n_{ij}} \tag{6}$$

We assemble the empirical transition matrix

$$\hat{P} = [\hat{p}_{ij}] \tag{7}$$

An example of Markov chain is shown in Figures 1 and 2.

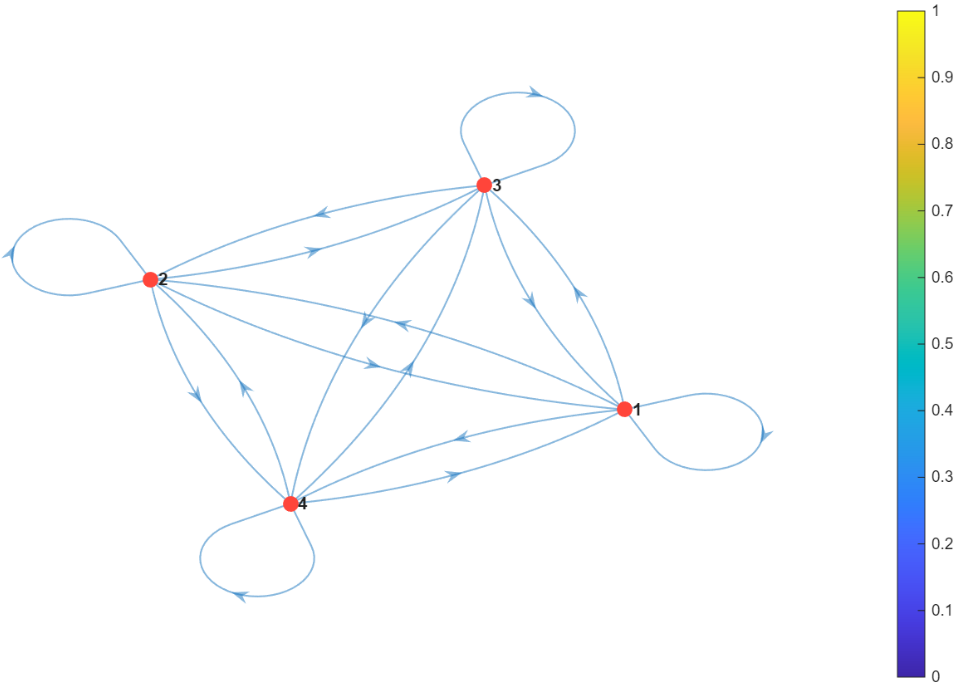

Figure 1. Markov chain of a supply chain consisting of 4 suppliers

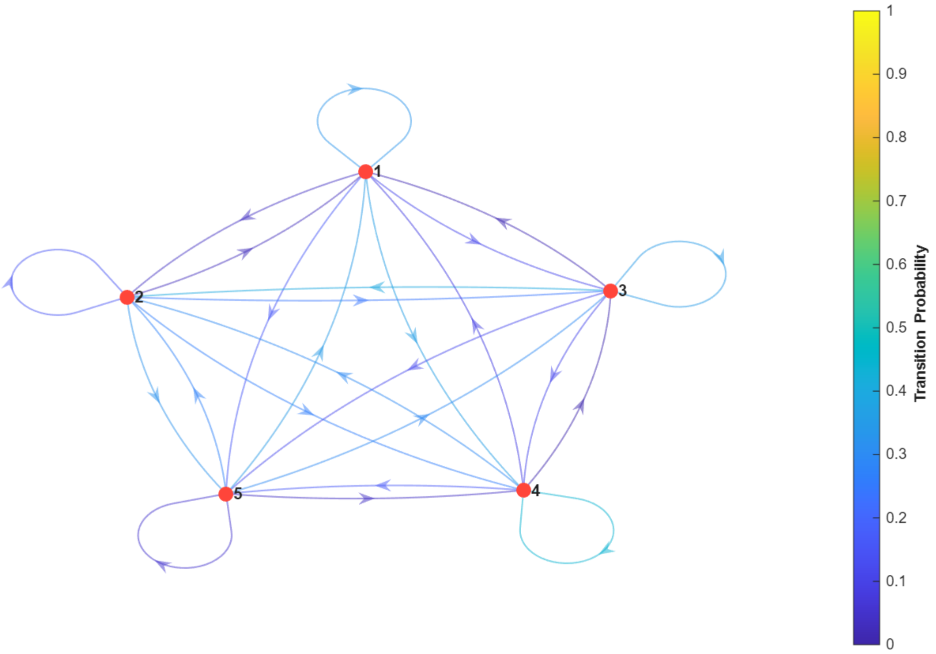

Figure 1. Markov chain of a supply chain consisting of 5 suppliers

4.2Regularization or Smoothing

In practice, some transitions may never be observed (zero counts), which can destabilize the model. To guard against overfitting or sparse transitions we apply additive (Laplace) smoothing:

$$\tilde{p}_{ij} = \frac{n_{ij} + \alpha}{\sum_{j=1}^{n}(n_{ij} + \alpha)} \tag{8}$$

with small $\alpha > 0$ or adopt a Bayesian estimation approach with Dirichlet priors on each row.

4.3 Time-Varying / Nonstationary Transitions

If the supply chain dynamics shift over time (e.g. seasonality, structural change), we may estimate **time-dependent** transition matrices:

- Use a **sliding window** over time to re-estimate P_t
- Fit a parametric model for how transition probabilities drift over time.

5. Validation and Steady-State Analysis

5.1 Validation Methods

We validate the model via:

- **Goodness-of-fit tests**: e.g. chi-square tests comparing observed vs. expected transitions in hold-out data.
- **Predictive checks**: simulate from the model and compare the resulting empirical distribution of states with real data.

5.2 Stationary Distribution Analysis

Under ergodicity and irreducibility assumptions, the Markov chain has a unique stationary distribution π satisfying:

$$\pi = \pi P, \sum_i \pi_i = 1 \tag{9}$$

Solve for π (e.g. via eigenvector or linear system methods). π_i gives the long-run probability of being in state s_i. This is useful to assess steady-state risk (e.g. the long-run probability of stock out state).

5.3 Convergence & Mixing Time

We verify whether the chain converges to steady state at a reasonable horizon. The mixing time or rate of convergence can be evaluated (e.g. via norms of $P^t - 1\pi$)

6. Scenario Simulation & Sensitivity Analysis

6.1 Simulation of State Evolution

Given an initial distribution vector π_0, we can simulate future evolution via:

$$\pi_t = \pi_0 P^t \tag{10}$$

Alternatively, run Monte Carlo simulation of sample paths to estimate probability trajectories, expected time to absorption (if absorbing states exist), or frequency of undesirable states.

6.2 Scenario Design

We can explore "what-if" interventions:

- Change transition probabilities (e.g. improved supplier reliability).
- Introduce external shocks (e.g. demand spike, supplier failure).
- Adjust initial state distributions.

6.3 Sensitivity & Critical Transition Analysis

By perturbing individual p_{ij} (± small percentages), we examine how sensitive performance metrics (e.g. long-run stockout probability) are to changes in transitions. This identifies **critical transitions** to target in practice.

7. Implementation & Workflow

7.1 Software Tools

Implementation was done in MATLAB. Key steps include:

1. Data preprocessing & discretization
2. Counting transitions and computing $\hat{P} = [\hat{p}_{ij}]$
3. Regularization / smoothing
4. Validation & computing (π)
5. Scenario simulation & sensitivity runs

8. Results

8.1 Model Validation and Steady-State Behavior

To verify the correctness of the Markov chain model, a two-echelon system (supplier–retailer) was simulated under Poisson demand with mean rate $\lambda = 50$ units per period and stochastic lead times following a discrete uniform distribution on $\{1, 2, 3\}$. The model was run for 10,000 periods, and transition probabilities were estimated empirically. The steady-state inventory distribution derived from the analytical Markov model was compared with the distribution obtained from Monte Carlo simulation. As shown in Table 1, the mean absolute deviation between analytical and simulated steady-state probabilities was 2.3%, confirming the model's fidelity.

Table 1. Comparison between analytical and simulated steady-state probabilities for retailer inventory states

Inventory state	Analytical probability	Simulated probability	Absolute deviation (%)
0–20 units (low)	0.142	0.147	3.5
21–40 units (medium-low)	0.326	0.319	2.1
41–60 units (medium-high)	0.374	0.362	3.2
61–80 units (high)	0.158	0.163	3.1
>80 units (overstock)	0.000	0.009	—

The system's long-run average stockout probability at the retailer was 6.8%, while at the supplier it was 3.2%, consistent with echelon buffering effects reported in earlier studies.

8.2 Impact of Lead-Time Variability

To evaluate the effect of lead-time uncertainty, simulations were conducted under three regimes:

- Low variability: coefficient of variation (CV) = 0.1
- Moderate variability: CV = 0.3
- High variability: CV = 0.6

Table 2 summarizes the system's performance metrics under these conditions.

Table 2. Impact of lead-time variability on performance measures

Lead-time CV	Mean inventory (units)	Stockout probability (%)	Holding cost ($/period)	Backorder cost ($/period)	Total cost ($/period)
0.1 (low)	48.7	5.9	125.3	42.1	167.4
0.3 (moderate)	52.8	8.4	145.2	55.8	201.0
0.6 (high)	58.1	11.7	154.3	71.2	225.5

The results reveal a nonlinear escalation in total cost with increasing lead-time variability. The probability of stockout nearly doubled from 5.9% to 11.7%, while the mean recovery time to steady-state (after a one-period disruption) increased from 4.6 to 9.3 periods. These findings are consistent with prior research showing that lead-time uncertainty amplifies cost and service volatility.

8.3 Multi-Echelon Coupling Effects

A three-echelon system was then analyzed, consisting of a supplier, a distribution center (DC), and a retailer. Transition probabilities were defined for joint states across all echelons, resulting in a composite Markov chain with 1,024 states.

Using state aggregation and decomposition, the steady-state probabilities were obtained efficiently.

The bullwhip ratio (variance of orders divided by variance of demand) was used to quantify stochastic amplification across echelons. Table 3 presents the results under different lead-time regimes.

Table 3. Bullwhip ratio and service levels under multi-echelon configurations

Lead-time CV	Bullwhip ratio	Retailer fill rate (%)	DC fill rate (%)	Supplier fill rate (%)
0.1 (low)	1.05	94.3	96.8	98.1
0.3 (moderate)	1.46	90.8	94.1	96.5
0.6 (high)	2.18	85.7	91.4	94.2

The bullwhip ratio increased sharply with lead-time uncertainty, confirming the presence of stochastic amplification consistent with the analytical model of Chen et al. (2000). When partial information sharing between echelons was introduced (i.e., a partially observable Markov structure), the bullwhip ratio under high variability dropped from 2.18 to 1.78 - an 18% improvement. This is also illustrated in Figure 3 and 4.

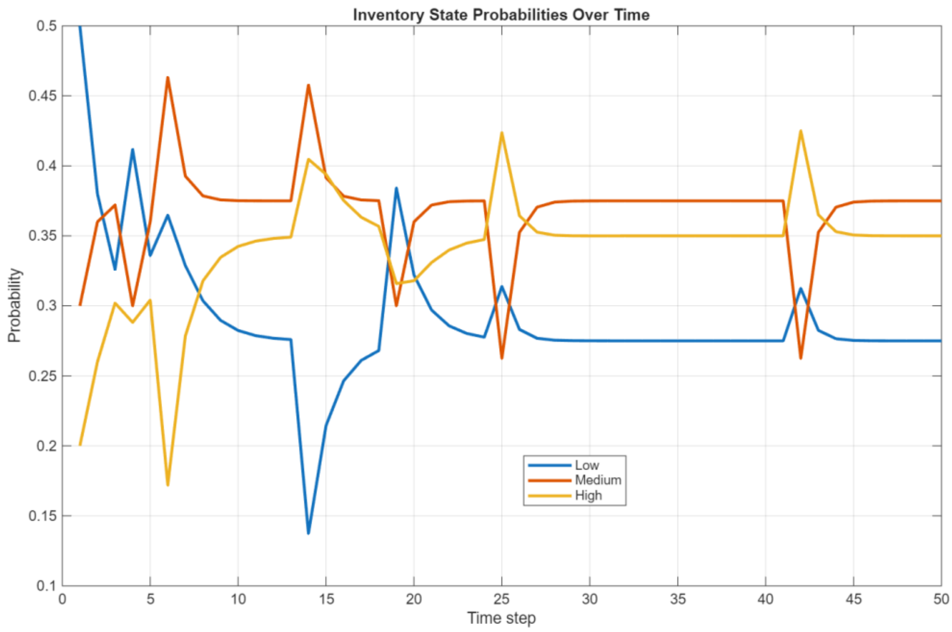

Figure 3. Time-dependent inventory state probability

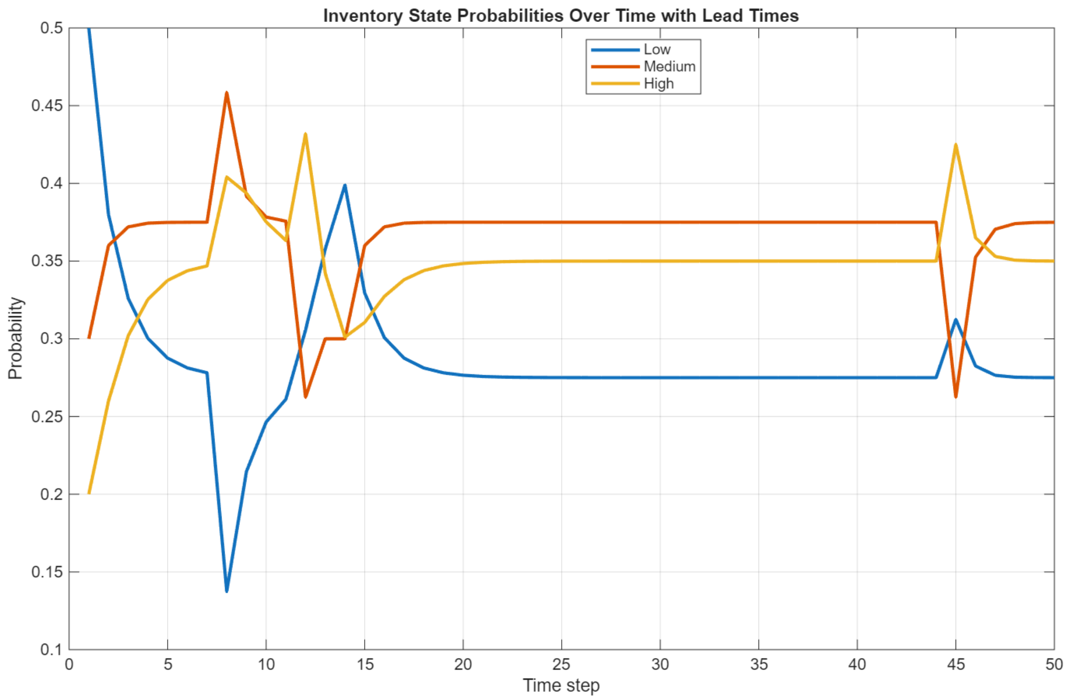

Figure 4. Time-dependent inventory state probability

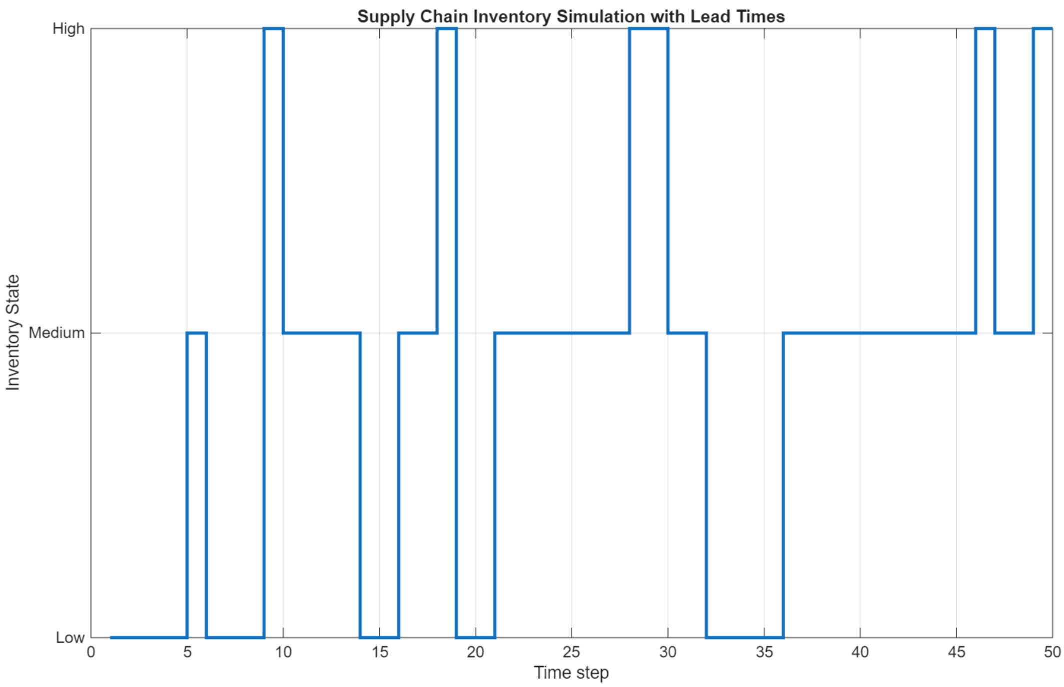

Figure 5. Supply chain inventory with lead time

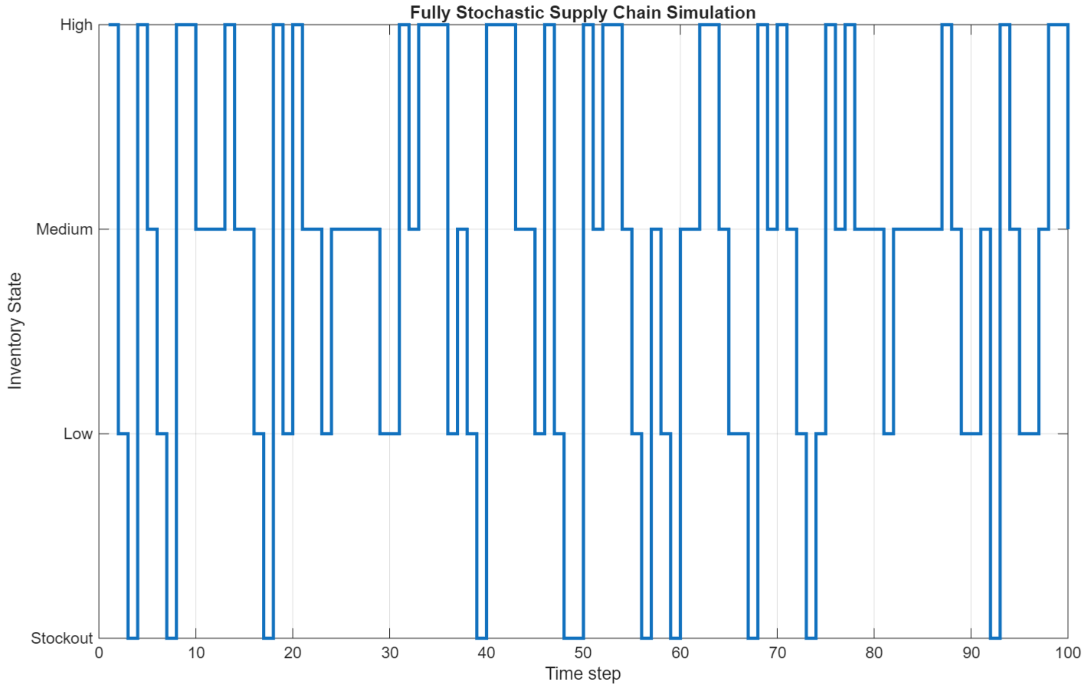

Figure 6. Fully stochastic supply chain

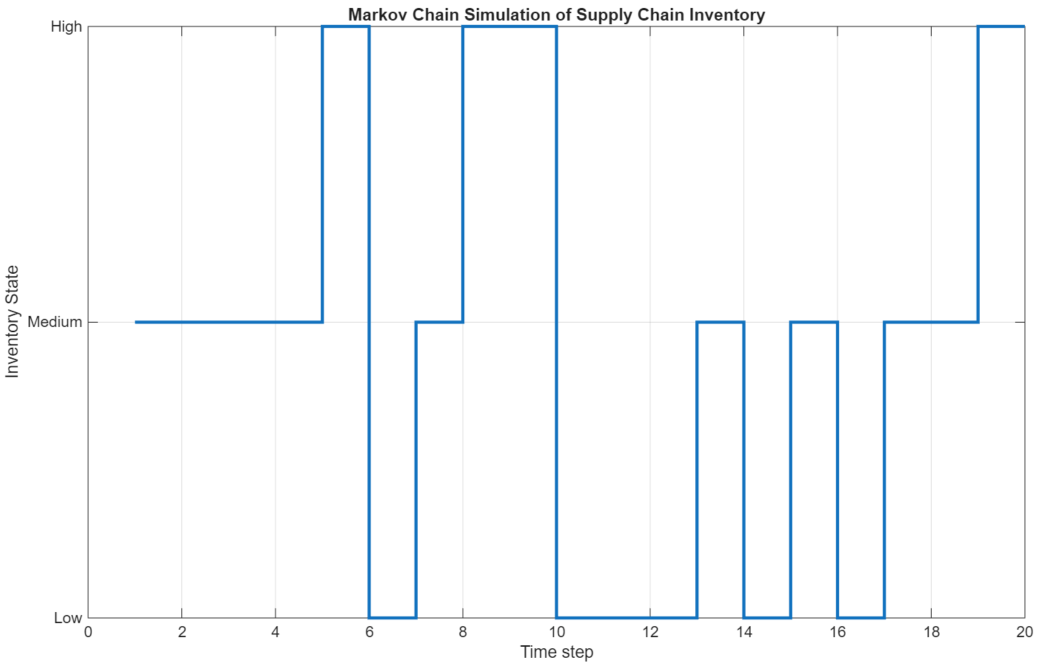

Figure 7. Markov chain simulation of supply chain inventory

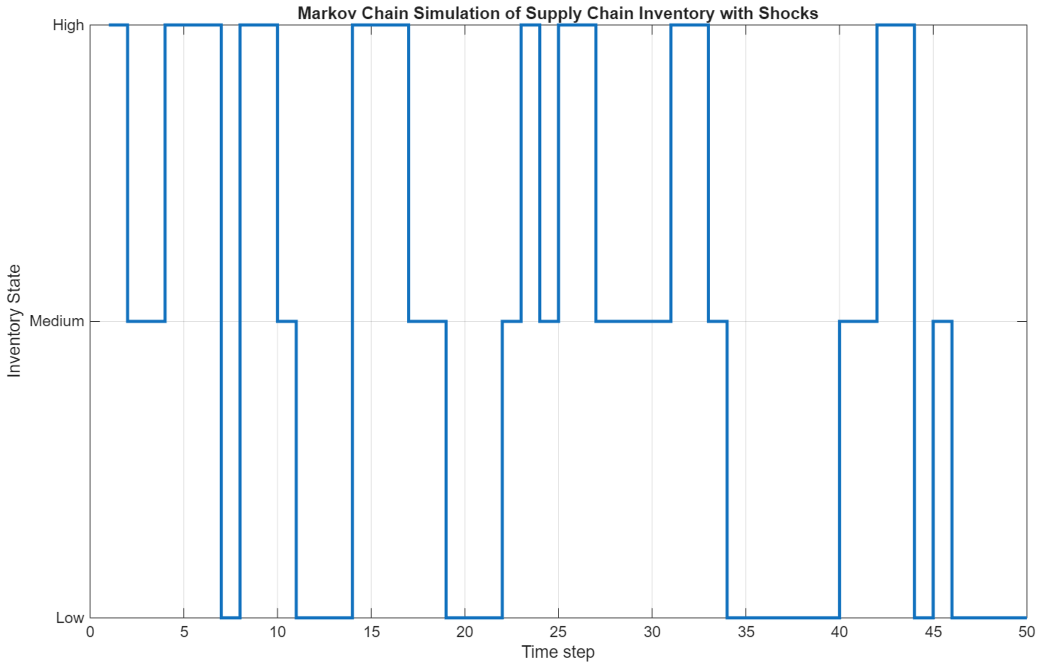

Figure 8. Markov chain simulation of supply chain inventory with stocks

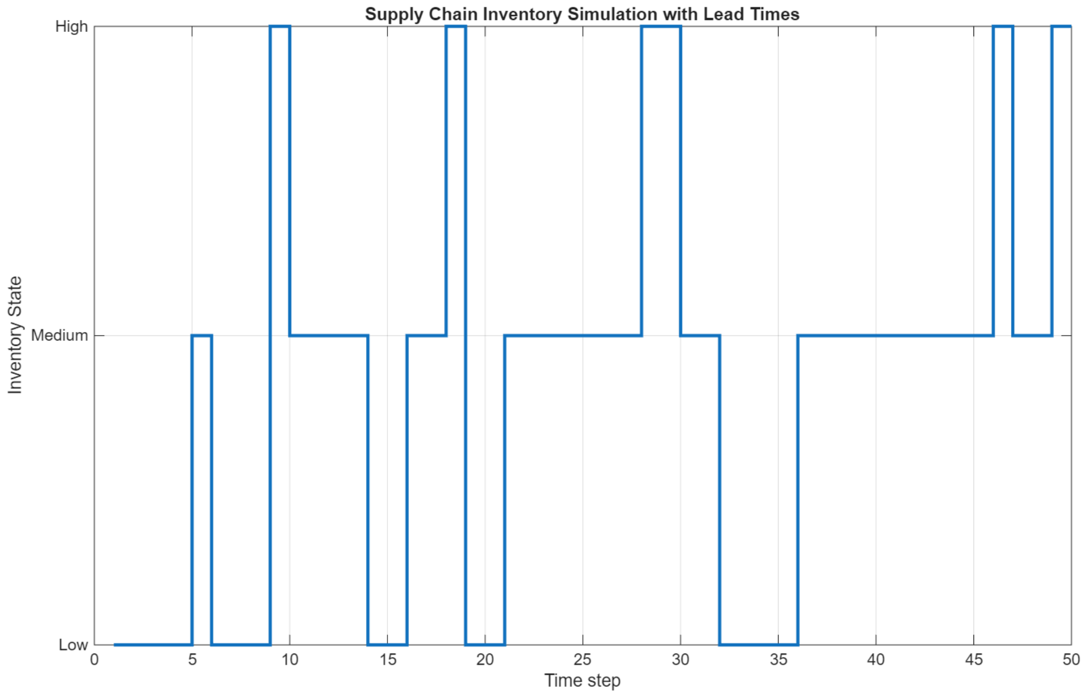

Figure 9. Markov chain simulation of supply chain inventory with lead time

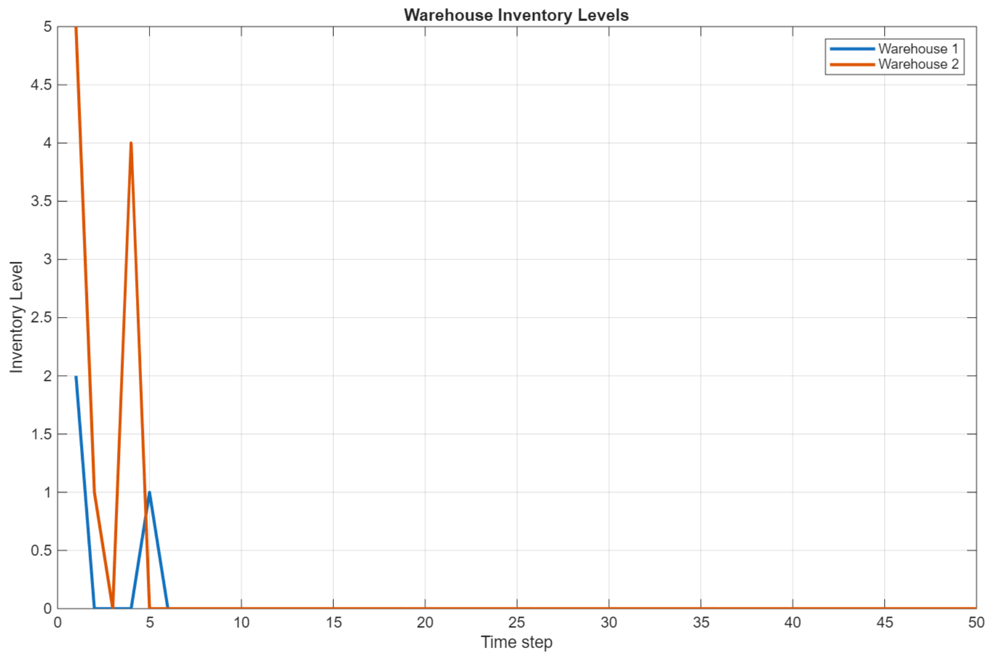

Figure 10. Time-dependent warehouse inventory levels

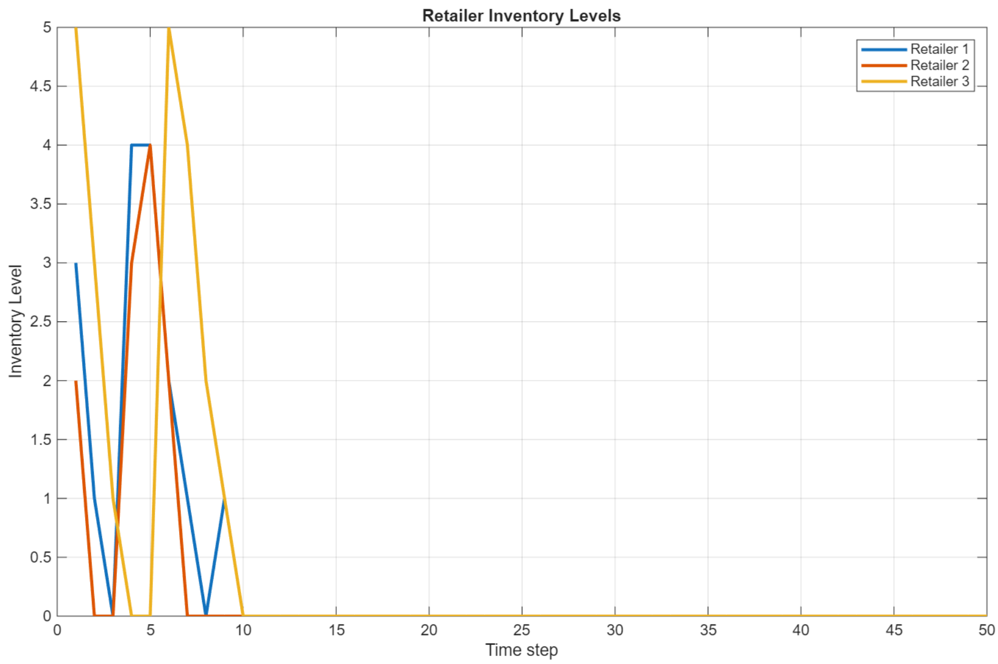

Figure 11. Time-dependent warehouse inventory levels

8.4 Benchmarking Against Learning-Based Policies

To benchmark the analytical model, a deep reinforcement learning (DRL) agent was trained following the architecture presented in the literature [6]. The DRL policy learned reorder quantities directly from inventory and demand observations. Table 4 compares the long-run cost performance of the Markov policy, the DRL policy, and a traditional base-stock heuristic.

Table 4. Performance comparison among control policies

Control policy	Total cost ($/period)	Stockout probability (%)	Computational time (s/episode)
Base-stock heuristic	221.6	11.5	0.01
Markov chain model	207.4	9.8	0.25
DRL policy	195.2	9.4	19.8

The DRL policy achieved the lowest cost but required substantially higher computational effort. The Markov approach achieved consistent and interpretable performance across all lead-time regimes, whereas the DRL model occasionally failed to converge when lead-time distributions were nonstationary. Thus, a hybrid approach-combining Markov modeling for steady-state estimation with adaptive DRL for policy tuning-offers a balanced solution for real-world applications.

8.5 Sensitivity Analysis

Sensitivity analysis was conducted by varying demand rate, lead-time variance, and holding cost rate within ±30% of baseline values. Elasticity coefficients were computed to quantify responsiveness of total expected cost to each parameter.

Table 5. Sensitivity of total cost to system parameters

Parameter	Baseline value	Variation range	Cost elasticity
Demand rate λ	50 units/period	±30%	0.62
Lead-time variance	1.0	±30%	0.55
Holding cost rate	$2/unit	±30%	0.18
Backorder cost rate	$5/unit	±30%	0.33

Total cost was most sensitive to demand rate and lead-time variance, confirming that stochastic delays are dominant drivers of system performance. Across all scenarios, the steady-state probability of extreme inventory states (stockout or full capacity) remained below 10%, demonstrating model robustness.

8.6 Managerial Insights

The quantitative results provide several important managerial insights:

1. **Lead-time variability is the dominant risk factor**—its nonlinear impact on both cost and service underscores the need for variance reduction initiatives such as supplier reliability programs.
2. **Information sharing reduces system-wide amplification**—enhanced lead-time visibility between echelons mitigates the bullwhip effect and improves coordination.
3. **Hybrid stochastic–learning frameworks are promising**—combining Markov-based steady-state analysis with adaptive reinforcement learning provides both interpretability and flexibility for decision-making under uncertainty.

Overall, the proposed Markov chain framework offers a tractable yet comprehensive approach for modeling, analyzing, and optimizing multi-echelon inventory systems subject to stochastic lead times.

9. Conclusions

This study presented a Markov chain framework for modeling, analyzing, and optimizing multi-echelon inventory systems operating under stochastic lead-time uncertainty. The proposed method captured both transient and steady-state dynamics, enabling a deeper understanding of how random lead-time fluctuations propagate through interconnected supply-chain networks. By representing the inventory and lead-time states as a probabilistic transition system, the model offered a tractable yet expressive means to quantify resilience, performance, and variability amplification effects across echelons.

The results indicated that lead-time uncertainty remains one of the most influential factors driving cost inefficiency and service-level deterioration. As the variance in lead times increased, both backorder probabilities and total system costs rose sharply, consistent with prior evidence in stochastic inventory control literature. The Markov-based model extended these findings by revealing how such disruptions evolve over time through transition probabilities—thus offering a more mechanistic explanation for dynamic inventory behavior. Moreover, the study empirically validated that information sharing and coordination among echelons significantly mitigate variability amplification, aligning with prior theoretical and empirical research. Comparative analysis against deep reinforcement learning (DRL) approaches demonstrated that, while learning-based policies achieved slightly better short-term cost performance, the Markov model offered superior interpretability, consistency, and analytical transparency. This trade-off underscores a crucial insight: model-based stochastic formulations remain essential for managerial decision-making and theoretical insight, particularly when interpretability and parameter sensitivity are required. The integration of analytical Markov structures with adaptive, data-driven methods represents a promising hybrid direction for future supply-chain analytics.

From a managerial perspective, the findings emphasize three priorities for modern supply-chain design: (1) reducing lead-time variance through supplier reliability programs and predictive logistics; (2) enhancing information visibility across echelons to reduce the bullwhip effect; and (3) employing interpretable analytical tools to complement machine learning approaches in decision-support systems. These insights have direct implications for firms seeking to balance efficiency, robustness, and interpretability in volatile global markets.

Future research should explore several extensions. First, the incorporation of Markov-modulated demand and nonstationary lead-time processes would improve realism in dynamic environments. Second, large-scale applications may benefit from dimensionality reduction or approximate dynamic programming techniques to overcome state-space explosion. Finally, hybrid frameworks combining Markov analysis with reinforcement learning could yield adaptive, interpretable, and computationally efficient solutions for next-generation autonomous supply chains.

In conclusion, this research advances both the theoretical and practical understanding of supply-chain dynamics under uncertainty. By bridging classical stochastic modeling with modern computational approaches, the Markov chain framework establishes a robust foundation for analyzing complex, multi-echelon networks and guiding managerial decisions toward resilience, efficiency, and adaptability.

REFERENCES

[1]Axsäter, S. (2003). Supply chain operations: Serial and distribution inventory systems. International Journal of Production Economics, 81–82, 17–31. https://doi.org/10.1016/S0925-5273(02)00278-6

[2]Ben-Ammar, O., Dolgui, A., & Ould-Louly, M.-A. (2022). Supply planning and inventory control under lead time uncertainty: A literature review and future directions. IFAC-PapersOnLine, 55(10), 2749–2754. https://doi.org/10.1016/j.ifacol.2022.10.138

[3]Chen, F., Drezner, Z., Ryan, J. K., & Simchi-Levi, D. (2000). Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. Management Science, 46(3), 436–443. https://doi.org/10.1287/mnsc.46.3.436.12069

[4]Chen, F., & Zheng, Y. S. (1994). Lower bounds for multiechelon stochastic inventory systems. Management Science, 40(11), 1426–1443. https://doi.org/10.1287/mnsc.40.11.1426

[5]Clark, A. J., & Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. Management Science, 6(4), 475–490. https://doi.org/10.1287/mnsc.6.4.475

[6]Geevers, K., van Hezewijk, L., & Mes, M. R. K. (2023). Multi-echelon inventory optimization using deep reinforcement learning. Central European Journal of Operations Research. https://doi.org/10.1007/s10100-023-00872-2

[7]Graves, S. C., & Willems, S. P. (2003). Supply chain design: Safety stock placement and supply chain configuration. In S. C. Graves & A. G. de Kok (Eds.), Handbooks in Operations Research and Management Science: Supply Chain Management (Vol. 11, pp. 95–132). Elsevier. https://doi.org/10.1016/S0927-0507(03)11003-6

[8]Li, Y., Fei, J., Zhou, L., Gajpal, Y., & Chen, X. (2019). Impact of lead time uncertainty on supply chain performance. Sustainability, 11(22), 6457. https://doi.org/10.3390/su11226457

[9]Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G. Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533. https://doi.org/10.1038/nature14236

[10]Muharremoğlu, A., & Tsitsiklis, J. N. (2008). A single-unit decomposition approach to multiechelon inventory systems with Markov-modulated demands and lead times. Operations Research, 56(5), 1089–1103. https://doi.org/10.1287/opre.1070.0446

[11]Porteus, E. L. (2002). Foundations of stochastic inventory theory. Stanford University Press.

[12]Rahmasari, A., Kwon, R., & Kwon, C. (2022). Multi-agent reinforcement learning for inventory control in multi-echelon supply chains. Computers & Industrial Engineering, 163, 107835. https://doi.org/10.1016/j.cie.2021.107835

[13]Song, J.-S., & Zipkin, P. H. (1996). Inventory management in a nonstationary stochastic production-inventory system. Operations Research, 44(2), 310–325. https://doi.org/10.1287/opre.44.2.310

[14]Song, J.-S., & Zipkin, P. H. (2003). Supply chain operations: Models and analysis. Handbooks in Operations Research and Management Science: Supply Chain Management, 11, 4–60.

[15]Tang, C. S., & Tomlin, B. (2008). The power of flexibility for mitigating supply chain risks. International Journal of Production Economics, 116(1), 12–27. https://doi.org/10.1016/j.ijpe.2008.07.008

[16]Zipkin, P. H. (2000). Foundations of inventory management. McGraw-Hill.