

Original Article

Reinforcement Learning for Dynamic Quality Optimization in Inventory Flow

NALINISIVAKUMAR

Department of physics, Kunthavai Naachchiyar Govt arts and science college Thanjavur.

ABSTRACT: *Efficient inventory flow management requires balancing cost, service level, and product quality in environments characterized by uncertainty and dynamic demand. Traditional inventory optimization models often rely on static assumptions and predefined decision rules, limiting their ability to adapt to real-time changes in quality degradation, demand variability, and operational constraints. This paper proposes a reinforcement learning-based framework for dynamic quality optimization in inventory flow systems. By modeling inventory decisions as a sequential decision-making problem, the proposed approach enables an intelligent agent to learn optimal ordering, holding, and allocation policies that minimize quality loss while maintaining operational efficiency. The framework incorporates quality-aware reward functions and state representations that capture inventory levels, product age, demand uncertainty, and quality decay dynamics. Experimental results demonstrate that reinforcement learning policies outperform conventional inventory control models in reducing quality-related losses, improving service levels, and adapting to fluctuating operating conditions. The findings highlight the potential of reinforcement learning as a powerful tool for achieving dynamic, data-driven inventory quality optimization in modern supply chain systems.*

KEYWORDS: *Reinforcement Learning, Inventory Flow Optimization, Quality Management, Dynamic Decision-Making, Supply Chain Analytics, Quality Degradation, Intelligent Inventory Systems, Operations Research.*

1. INTRODUCTION

1.1. BACKGROUND OF INVENTORY FLOW AND QUALITY MANAGEMENT CHALLENGES

At the core of supply chain operations is inventory flow management, which dictates how products are moved between all stages of storage, handling and distribution in order to provide customers with the products they desire. For example, in many real-world systems that concern products whose quality deteriorates over time (e.g. perishable or fragile products), inventory decisions directly impact the quality of the product. There is a risk that they will degrade in quality (spoil) from aging, prolonged storage uptake environmental and operational conditions or delays in full operations causing spoilage, obsolescence and customer dissatisfaction. These challenges continue to be exacerbated by increasing market volatility, an uncertain demand environment and tighter delivery expectations. With supply chains evolving to be more complex, dynamic and the need to maintain an optimal inventory-flow efficiency vs quality retention balance has become a pivotal managerial challenge.

1.2. LIMITATIONS OF TRADITIONAL INVENTORY AND QUALITY OPTIMIZATION MODELS

Conventional inventory and quality optimization models are mainly driven by deterministic or stochastic optimization frameworks with pre-established parameters and static decision rules. Classic models usually rely on known demand distributions, constant lead times, and simplified quality decay functions that rarely hold true in practice. Stochastic extensions bring uncertainty to these models but still require the specification of probability distributions and system dynamics. Such assumptions hampers their ability to respond dynamically as demand fluctuates in real time, or that supply may be compromised at some point due to a restrictive change or quality situation. Furthermore, these models are usually short-term optimizers and are incapable of capturing the long-term effect on quality deterioration in inventory decisions leading to a either a suboptimal or local solution especially in dynamic environments.

1.3. MOTIVATION FOR REINFORCEMENT LEARNING IN DYNAMIC DECISION-MAKING

It forms a data driven and adaptive strategy to choose out actions in uncertain, delayed reward and complex system interactive environments. Instead of requiring an a priori model of the dynamics of the system to be optimized for, reinforcement learning learns optimal policies through trial and error while interacting with the environment. This characteristic makes it particularly well-suited for inventory systems where the demand pattern, quality degradation rate, and operational conditions evolve over time. Framing inventory flow management as a sequential decision-making problem allows for using reinforcement learning in solving the question: The system can dynamically adjust decisions based on actual observed outcomes, by trading off immediate operational costs against longer-term preservation of quality.

1.4. RESEARCH OBJECTIVES AND CONTRIBUTIONS

The major purpose of this study is to provide a framework for reinforcement learning–based dynamic quality optimization in inventory flow systems. We model inventory decision as a sequential learning problem that directly incorporates quality degradation in the optimization of inventory. The major contributions include: (1) the proposed a reinforcement learning model for optimizing the quality-aware inventory flow; (2) coupling inventory dynamics with a quality decay mechanism into an effective learning environment, so that many important features can be more effectively integrated, and finally, (3) providing extensive comparisons to evaluate the proposed framework with existing inventory optimization methods. This provides an advancement in the literature by cross-linking reinforcement learning with quality-cognizant inventory management and a real-world methodology for smart supply chain decision making.

2. LITERATURE REVIEW

2.1. CLASSICAL INVENTORY CONTROL AND QUALITY OPTIMIZATION MODELS

Traditional inventory control models like the Economic Order Quantity and reorder point are based either on minimizing total costs associated with inventory under a simplified assumption. Extensions of these models include quality considerations, for example by incorporating spoilage/together with deterioration and defect rate into the model. These quality-aware models enhance realism, but they are relatively static and rely heavily on manually estimating the model parameters. A limitation of their analytical nature is that they cannot adapt to changes in system dynamics and thus malfunction, especially during cases with high degrees of uncertainty.

2.2. STOCHASTIC AND DYNAMIC INVENTORY FLOW MANAGEMENT

Stochastic and dynamic inventory models generalize classical methods by introducing random demand, lead times and quality decay. Stochastic programming or Markov decision process is frequently used in these models which better incorporates the uncertainty. These approaches are theoretically sound, but in practice require exact knowledge of all transition probabilities and system parameters. In addition, when attempts are made to solve large-scale stochastic inventory problems through computational methods, the exponential increase in time complexity limits their practical use in situations relevant to real-time operational environments.

2.3. REINFORCEMENT LEARNING APPLICATIONS IN SUPPLY CHAIN AND OPERATIONS

Reinforcement learning has gained significant attention in recent years and is becoming increasingly popular in applications of supply chain and operations management. It is applied to inventory replenishment, production planning, logistics optimization, and other demand-responsive control systems by leveraging reinforcement learning. Experiments revealed that reinforcement learning can learn adaptive policies for uncertain states and therefore outperform traditional heuristics in these studies. Nevertheless, the majority of current applications pay little attention to directly modeling and optimizing the product quality evolution over time while they mainly focus on either minimizing costs or maximizing service levels.

2.4. IDENTIFIED RESEARCH GAPS

Although there has been a growing literature on RL in inventory management, substantial gaps remain. In particular, the explicit incorporation of quality degradation mechanisms in reinforcement learning formulations has been sparsely studied. Literature has either implicitly treated quality loss or blatantly neglected the elusive long-term effects on inventory performance. Moreover, little attention is paid to reward structures that jointly optimize inventory flow efficiency and quality preservation. These gaps provide the impetus of developing a reinforcement learning–based framework that appropriately tackle quality optimization in dynamic inventory systems.

3. PROBLEM FORMULATION AND THEORETICAL FRAMEWORK

3.1. INVENTORY FLOW DYNAMICS AND QUALITY DEGRADATION MECHANISMS

The inventory system in this study evolves over a time set, where the state of the inventory system changes due to ordering, demand fulfilment and holding decisions. Quality loss is modeled using an age-dependent quality deterioration function, where the product quality deteriorates as it remains in storage or as inventory ages. And under certain conditions, this degradation may happen more quickly, leading to spoilage or expiration. Therefore, capturing the interplay between inventory flow and quality decay is important because actions to minimize holding cost can lead a higher performance loss term while an aggressive control for the preservation of quality could increase operational costs.

3.2. STATE, ACTION, AND REWARD FORMULATION FOR REINFORCEMENT LEARNING

In the case of first reinforcement learning framework, states represent the quantity of items in inventory system at time t , actions — which item will be released from our warehouse to satisfy an existing order, reward — how to measure performance. State is essentially the current state of the system in as far as some features that characterize it, e.g. inventory levels, product age or quality distribution of products (for example, say for cosmetics) and demand-related information. Actions are operational decisions: these may be order quantities, allocation of inventory or removal of decayed items. The reward function states the trade-off between efficiency and quality — penalizing losing quality, high holding costs, and

stockout with timely fulfillment and service. This formulation gives the learning agent an opportunity to optimize for long-term performance and not for short-sighted gain.

3.3. CONSTRAINTS AND OPERATIONAL ASSUMPTIONS

Real world constraints such as limits on storage capacity, ordering thresholds and service level agreements are included in the problem formulation. These assumptions are made on the behavior of demand, lead times, and the processes of performance loss in order to provide tractable models which will learn reasonably stably. These assumptions are made to represent practical operating conditions while remaining as generalizable as possible. This step is followed by adding constraints to make certain that the learnt reinforcement learning policies are implementable within practical inventory systems.

3.4. THEORETICAL FOUNDATIONS OF RL-BASED OPTIMIZATION

The proposed approach is theoretically rooted in reinforcement learning and Markov decision process theory, which provides a formal framework for sequential optimization under uncertainty. Reinforcement learning is a type of algorithm that deals with learning an optimal policy based on an agent interacting with its environment to maximize the expected cumulative reward. The learning paradigm is particularly adapted to inventory systems, where decisions are delayed in their effect on producing costly and quality consequences that are also interdependent. The theoretical framework, built upon reinforcement learning theory provides a mechanism to adaptively perform data-driven optimization of inventory flow and quality management for supply chains.

4. METHODOLOGY

4.1. DATA SOURCES AND SIMULATION ENVIRONMENT

This study proposes a simulation-driven methodology that captures the dynamic behavior of inventory flow and quality degradation under uncertainty. A more realistic inventory may involve limited access to refined operational data in real world settings, thus we have created a controlled simulation gel to replicate plausible level of activity an observed level of activity. This simulation include stochastic dynamic demand, arrival lead times and dependent ageing quality processes. Industry-derived assumptions are used to calibrate input parameters, ensuring their practical relevance. Our environment progresses according to interleaved decision epochs, where the reinforcement learning agent continuously interacts with the system via state observations (inventories), actions (positions), and rewards. Such a setup allows for systematic testing in different demand and quality scenarios, while maintaining reproducibility and consistent policy assessment.

4.2. REINFORCEMENT LEARNING ALGORITHMS USED

Given that inventory decision-making is a sequential and stochastic process, this paper uses reinforcement learning (RL) algorithms to learn optimal policies from simulations of the environment. This is particularly suited to model-free reinforcement learning approaches, as it does not require knowledge on the system transition probabilities or quality decay dynamics. We explore both value-based and policy-based algorithms so that we can investigate the different aspects of decision-making. They are chosen due to their robustness in high-dimensional state spaces, the versatility under sparse rewards and long-term dependencies of inventory decisions through to quality outcomes. Learn reinforcement learning, so the policy could gradually change as we learn how to control inventory flow.

4.3. FEATURE REPRESENTATION AND STATE ENCODING

The proper features to have the RL agent understanding the inventory system. The state encoding is one of the most important elements in such formulations as it contains vital features about the system, including current on-hand inventories, age or quality distribution among items, observed demand in the recent past and remaining storage capacity. Features concerning quality are explicitly rooted to depict degradation dynamics as a function of time. Temporal information is introduced as regards to lagged observations, or using aggregated statistics so that the agent understands patterns in time and is able to predict what possible states of a system might be at a future point. This organized representation of state ensures a strongly contextualized scenario for the learning agent to optimally accomplish decision making by balancing inventory efficiency with quality preservation.

4.4. TRAINING, VALIDATION, AND POLICY EVALUATION STRATEGY

It iterates over many episodes of interaction between the reinforcement learning agent and a simulation environment where the combinations get gradually optimal decision policies. Set of actions that the agent tries to explore while training, to learn about long-term consequences on inventory cost and quality results. Generalization ability is validated by assessing the policies learnt on unseen demand and quality scenarios. Since the target policy is kept constant, policy evaluation naturally emphasizes long-run performance metrics over short-term rewards; any learned strategies are guaranteed to remain stable and policies functional under a varying operating context. Such a systematic training and evaluating approach makes the described methodology more robust and reliable.

5. REINFORCEMENT LEARNING–BASED QUALITY OPTIMIZATION FRAMEWORK

5.1. ARCHITECTURE OF THE PROPOSED RL FRAMEWORK

Our proposed reinforcement learning framework is developed in a modular architecture that incorporates inventory dynamics, quality degradation modeling and decision-making intelligence components. The framework contains an environment layer that replicates the inventory system, an agent layer for learning optimal policies and an evaluation layer that translates decisions into corresponding performance outcomes. The agent watches the state of the system, decides on inventory actions and gets rewards for being as cost-efficient and quality-preserving as possible. That means it is the continuous learning and adaptation of that framework through this closed-loop architecture that allows it to respond much better to changing demand patterns and quality conditions.

TABLE 1 Components of the Reinforcement Learning–Based Quality Optimization Framework

Component Layer	Description	Functional Role
Environment Layer	Represents the inventory system including demand patterns, inventory levels, and quality degradation dynamics	Simulates real-world inventory flow and provides state feedback
Agent Layer	Reinforcement learning agent responsible for policy learning and decision-making	Selects optimal inventory actions based on observed states
Action Module	Defines allowable inventory decisions such as ordering, allocation, and disposal	Enables control over inventory flow and quality outcomes
Reward Mechanism	Quantifies performance using cost efficiency and quality preservation measures	Guides learning toward long-term optimal behavior
Evaluation Layer	Assesses policy performance using inventory and quality metrics	Translates agent decisions into measurable outcomes

5.2. QUALITY-AWARE INVENTORY DECISION-MAKING PROCESS

Instead, awareness of quality is directly intertwined with the decision-making process of the reinforcement learning agent like a reward structure and state representation. Decisions concerning inventory are assessed both by a short-term view of costs to operations and the long-range benefit or harm to product quality. You optimize decisions about how much to order and where to store, as well as when to dispose of items that have gone bad avoiding incurring quality loss subject to required service levels. Ultimately, the framework can promote proactive decision making that more sustainably manages the flow of inventory by explicitly considering quality as less than full attenuation results in deterioration rather than disposal which shrinks waste streams.

5.3. EXPLORATION–EXPLOITATION STRATEGY

Reinforcement learning throughout recurring systems requires the exploration–exploitation trade-off, especially in inventory systems where wrong choices can result in high losses. In the early stages of training the agent must explore to learn what happens with an inventory action raised in different scenarios. In this rehearsal strategy over time it will change from exploration to non exploring, focusing solely on efficient policies learned. This avoids convergence to suboptimal recommendations too early, whilst allowing the agent eventually to settle on high-reward consistent decision rules. The exploration–exploitation mechanism allows the agent to cope with uncertainty and find high-quality optimization strategies.

5.4. INTEGRATION WITH INVENTORY PERFORMANCE METRICS

In order to test learned policies, our reinforcement learning framework is specified in conjunction with standard inventory performance metrics. Metrics like quality loss, holding cost, stockout frequency and service level are integrated into the reward and evaluation. The framework ensures that the generated reinforcement learning objectives are aligned with commonly accepted inventory performance measures, ensuring practical relevance and interpretability. This integration empowers decision-makers to evaluate the operational efficiency alongside quality outcomes of policy effectiveness which makes it suitable for adoption in practical inventory systems.

6. EXPERIMENTAL RESULTS AND ANALYSIS

6.1. POLICY PERFORMANCE EVALUATION

The reinforcement learning policies that they learn are then tested via large simulation experiments over a variety of demands and quality scenarios. The results show that the RL agent learns to adapt its inventory policy in response to the dynamic shifting of system conditions. The learned policies are stable in terms of lower-quality loss and better service levels, and improve over time. The evaluation demonstrates that reinforcement learning explores long term dependencies between inventory decisions and quality outcomes to achieve the best performance across the entire system.

6.2. COMPARISON WITH TRADITIONAL INVENTORY OPTIMIZATION MODELS

This paper compares the reinforcement learning-based approach with conventional inventory optimization models implemented in practice. Traditional models generally assume fixed reorder points or static optimization regulations that do not update emphatically when the underlying dynamics of the system change. Our experiments illustrate that the reinforcement learning framework consistently excels over these conventional methods, especially within environments featuring high variability and decayed quality. Reinforcement Learning, being adaptive allows dynamic and flexible inventory control to minimize inefficiencies that would otherwise occur with the static decision rules.

6.3. IMPACT ON QUALITY LOSS, HOLDING COST, AND SERVICE LEVEL

The experimental results show that the reinforcement learning framework can effectively balance quality, cost, and service. In contrast to traditional models, perhaps the most significant reduction in the loss of quality is achieved because Proactive Capacity Management is employed to control inventory age and flow. While maintaining or improving service levels through better demand responsiveness, more accurate order decisions will lead to optimal holding costs! These results confirm that reinforcement learning allows for the optimal multi-objective treatment of inventories, which may be in conflict.

6.4. SENSITIVITY ANALYSIS UNDER DEMAND AND QUALITY UNCERTAINTY

The robustness of the reinforcement learning policies is examined through a sensitivity analysis under different demand unknowability scenarios and levels of quality degradation. The results illustrate that the policies learned by the proposed method are robust to increasing system uncertainty. Second, the agent alters its ordering and allocation decisions to reduce risk exposure to increased variability. This even more robustness illustrates that reinforcement learning is a good fit for real world inventory systems where uncertainty and variability will be present.

7. DISCUSSION

7.1. INTERPRETATION OF RESULTS

Our experimental results show that reinforcement learning serves as a very effective vehicle for simultaneous optimisation of inventory flows & quality outcomes within dynamic and uncertain environments. The better performance of the learned policies is due to their internalization of long-term system behavior instead of short term heuristics. Through the ongoing observation of system states over time and a method of learning by observing delayed rewards, the reinforcement learning agent captures intricate interdependencies between inventory ageing with quality degradation, demand variability and cumulative costs. The results show that quality-aware decision-making can significantly lower waste and obsolescence levels while keeping service rates unchanged. This result highlights the value of incorporating quality dynamics directly into inventory control models, since neglecting degradation effects leads to inefficient and reactive policies.

7.2. MANAGERIAL AND OPERATIONAL IMPLICATIONS

The results imply a more informed and flexible operational decision-making context provided by reinforcement learning-based inventory systems from the manager's perspective. Data-driven policies that dynamically respond to changing demand and quality conditions allow managers to go beyond static reorder rules. Since it reduces the loss in quality and costs of holding such inventories, its economic advantages are particularly evident in industries where perishable, time-sensitive or quality-sensitive products (such as foodstuffs, pharmaceuticals and chemicals) are handled. From an operational standpoint, the framework facilitates proactive actions like adjusting replenishment proactively or disposing of inventory when needed to better support supply chain resilience. Performance metrics are also interpretable, allowing managers to directly relate reinforcement learning outcomes with business objectives and commitments on levels of service.

7.3. SCALABILITY AND REAL-WORLD APPLICABILITY

Aiming for scalability, the proposed reinforcement learning framework can be extended to larger inventory system. It has a modular architecture that works with existing enterprise resource planning and inventory management systems. Currently only implemented in a simulation environment, this learning can be applied to the real world using historical data as well as live online incremental updates. The feasibility of implementing reinforcement learning at scale is further enforced by the advances in computational infrastructure and cloud based deployment. These attributes suggest significant promise for real-life application — especially in a context where organisations are ever more in need of intelligent automation to react to complex and uncertain operational landscapes.

8. LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

8.1. MODEL ASSUMPTIONS AND DATA LIMITATIONS

Though the proposed approach shows promise, it relies on a number of modeling assumptions that may hinder its generality. The simulation environment features fixed demand distributions and known quality degradation functions, which may not represent all real-world complexities. Furthermore, the quality deterioration process is modeled in a more rigid structure which may vary across industries or product categories. Simulated data is needed to perform well defined experiments on entire neural network architectures, but you need large quantities of real-world data where this cannot be feasibly performed. Future research should include empirical data and more generalised patterns of quality degradation for a realistic application.

8.2. COMPUTATIONAL COMPLEXITY AND TRAINING CHALLENGES

While training reinforcement learning algorithms typically require a lot of time since they use up lots of computer resources, the problems become exponentially larger as the state and action space grows. Real-world inventory systems could benefit from appropriate safeguards since the exploration phase inside the learning process can be very expensive or infeasibly large. In addition hyperparameter tuning and convergence-stability provide even more challenges as well. These computational considerations may limit the immediate deployability of these methods in resource-constrained environments. We can explore the more efficient algorithms, approximations and transfer learning in the future research to help an agent learn faster while laying less overhead.

TABLE 2 Computational Complexity and Training Challenges

Challenge Area	Description	Impact on Implementation	Potential Mitigation
State Space Size	High-dimensional inventory and quality states	Increased memory and computation cost	State aggregation and dimensionality reduction
Action Space Expansion	Multiple inventory control actions	Slower convergence	Hierarchical or parameterized actions
Training Time	Long learning horizons	Delayed policy deployment	Transfer learning and warm-start policies
Hyperparameter Sensitivity	Learning rate and exploration settings	Unstable performance	Automated tuning and adaptive learning rates

8.3. EXTENSIONS TO MULTI-ECHELON AND REAL-TIME SYSTEMS

Extending the proposed framework to multi-echelon inventory systems (with Suppliers, Warehouses and Retailers) will be an important avenue of future research. This adds to the complexity by creating interdependencies across multiple decision layers and introducing information delays. The incorporation of live real-time data streams from sensors, Internet of Things devices and demand forecasting systems can also increase the accuracy, speed response time to market dynamics. Furthermore, decentralized or multi-agent reinforcement learning frameworks can be investigated to help collaborative decision-making between supply chain partners. These would further extend the reach and influence of centralized-reinforcement learning-based quality optimization in real-world inventory management applications.

9. CONCLUSION

This study proposed a reinforcement learning-based methodology for dynamic quality optimization of inventory flow systems, which is critical because traditional inventory control models are based on static assumptions and predefined decision rules. The proposed framework formulates inventory management as a sequential decision-making problem with quality-aware, adaptable policies that respond well to demand uncertainty, aging of inventory, and quality degradation dynamics. Experimental results show that reinforcement learning policies consistently outperform conventional inventory optimization methods when it comes to quality loss reduction, holding and disposal costs reduction, and service level maintenance in uncertain operating environments. The major result of this study is that explicitly modeling quality degradation in the learning process substantially improves long-term inventory performance because the agent learns to trade-off short-term operational efficiency with long-term quality preservation. The study extends the inventory and quality management literature through a unified framework that integrates reinforcement learning with quality-aware inventory decision-making, and furthers existing research by integrating two problems - inventory control, which is frequently treated as independent from quality management. On the other hand, from a practical perspective, the approach proposed shows a data driven decision-support mechanism with the ability to react in real time or quickly changing environments making it relevant for industries where perishability or time is an urgent product attribute. Additionally, the framework shows good potential for scalability and deployment in real-world scenarios with the usage of data from modern data infrastructures as well as enterprise systems. Although challenges remain for computing and data, the results suggest reinforcement learning represents a next-generation paradigm for intelligent inventory systems. The key takeaway here is that this research demonstrates the potential of reinforcement learning in becoming the basis and foundation for proactive, adaptive, and quality-based supply chain inventory management working towards re-inventing resilient smart factories supported during uncertain times.

REFERENCES

- [1] R. S. Sutton and A. Barto, *Reinforcement learning: An introduction*, 2nd ed. Cambridge, Ma ; London: The Mit Press, 2018.
- [2] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016, doi: <https://doi.org/10.1038/nature16961>.
- [3] "P. H. Zipkin, 'Foundations of Inventory Management,' McGraw-Hill Companies, New York, 2000. - References - Scientific Research Publishing," *Scirp.org*, 2025. <https://www.scirp.org/reference/referencespapers?referenceid=983941>
- [4] S. Nahmias and T. Olsen, *Production and operations analysis : strategy, quality, analytics, application*. Long Grove, Illinois: Waveland Press, Inc, 2015.

- [5] B. Van Roy, R. S. Sutton, S. Singh, and C. Szepesvári, “Foundations of reinforcement learning,” *Foundations and Trends in Machine Learning*, vol. 14, no. 3, pp. 197–346, 2021.
- [6] E. Mohebbi, and F. Choobineh, “Perishable inventory systems with multiple deterioration rates,” *European Journal of Operational Research*, vol. 165, no. 3, pp. 704–717, 2005.
- [7] I. Giannoccaro and P. Pontrandolfo, “Inventory management in supply chains: a reinforcement learning approach,” *International Journal of Production Economics*, vol. 78, no. 2, pp. 153–161, Jul. 2002, doi: [https://doi.org/10.1016/s0925-5273\(00\)00156-0](https://doi.org/10.1016/s0925-5273(00)00156-0).
- [8] D. P. Bertsekas, *Reinforcement learning and optimal control*. Belmont, Massachusetts: Athena Scientific, 2019.
- [9] A. G. Kök, M. L. Fisher, and R. Vaidyanathan, “Assortment Planning: Review of Literature and Industry Practice,” *Retail Supply Chain Management*, pp. 99–153, 2008, doi: https://doi.org/10.1007/978-0-387-78902-6_6.
- [10] R. N. Boute, S. M. Disney, M. R. Lambrecht, and B. Van Houdt, “A win–win solution for the bullwhip problem,” *International Journal of Production Economics*, vol. 108, no. 1–2, pp. 125–137, 2007.
- [11] W. B. Powell, *Approximate Dynamic Programming*. Wiley, 2011. doi: <https://doi.org/10.1002/9781118029176>.
- [12] C. S. Tang, “Robust strategies for mitigating supply chain disruptions,” *International Journal of Logistics Research and Applications*, vol. 9, no. 1, pp. 33–45, Mar. 2006, doi: <https://doi.org/10.1080/13675560500405584>.
- [13] A. Pandiammal, K. Iyna, “The Causality between FII and BSE in India,” *International Journal of Applied Social Science*, vol. 6, no. 6, pp. 1357-1361, 2019.