processes

MDPI

Article

Algorithmic Approaches to Inventory Management Optimization

Hector D. Perez, **Christian D. Hubbs**, **Can Li** and **Ignacio E. Grossmann** *

Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15123, USA; hperezpa@andrew.cmu.edu (H.D.P.); chubbs@andrew.cmu.edu (C.D.H.); canl1@andrew.cmu.edu (C.L.)
* Correspondence: grossmann@cmu.edu; Tel.: +1-412-268-3642

Abstract: An inventory management problem is addressed for a make-to-order supply chain that has inventory holding and/or manufacturing locations at each node. The lead times between nodes and production capacity limits are heterogeneous across the network. This study focuses on a single product, a multi-period centralized system in which a retailer is subject to an uncertain stationary consumer demand at each time period. Two sales scenarios are considered for any unfulfilled demand: backlogging or lost sales. The daily inventory replenishment requests from immediate suppliers throughout the network are modeled and optimized using three different approaches: (1) deterministic linear programming, (2) multi-stage stochastic linear programming, and (3) reinforcement learning. The performance of the three methods is compared and contrasted in terms of profit (reward), service level, and inventory profiles throughout the supply chain. The proposed optimization strategies are tested in a stochastic simulation environment that was built upon the open-source OR-Gym Python package. The results indicate that, of the three approaches, stochastic modeling yields the largest increase in profit, whereas reinforcement learning creates more balanced inventory policies that would potentially respond well to network disruptions. Furthermore, deterministic models perform well in determining dynamic reorder policies that are comparable to reinforcement learning in terms of their profitability.

Keywords: inventory management; supply chain; multi-echelon; stochastic programming; reinforcement learning

Citation: Perez, H.D.; Hubbs, C.D.; Li, C.; Grossmann, I.E. Algorithmic Approaches to Inventory Management Optimization. *Processes* **2021**, *9*, 102. https://doi.org/10.3390/pr9010102

Received: 1 December 2020
Accepted: 30 December 2020
Published: 6 January 2021

1. Introduction

Modern supply chains are complex systems that interconnect the globe. Efficient supply chains are able to control costs and ensure delivery to customers with minimal delays and interruptions. Inventory management is a key component in achieving these goals. Higher inventory levels allow for suppliers to maintain better customer service levels, but they come at a higher cost, which often gets passed on to their customers and, ultimately, to the end consumers. This is particularly the case for perishable items that have a limited shelf life and can go to waste if the inventory exceeds demand. Thus, every participant in the supply chain has an incentive to find the appropriate balance in inventory levels to maximize profitability and maintain market competitiveness. Efficient supply chains are able to coordinate material flows amongst its different stages to avoid the "bullwhip effect", whereby over corrections can lead to a cascading rise or fall in inventory, having a detrimental impact on the supply chain costs and performance [1].

Extensive literature exists in supply chain and inventory management. Relevant review papers in the area of inventory optimization include those of Eruguz et al. [2] and Simchi-Levi and Zhao [3]. The inventory management problem (IMP) that is presented in this work is built upon the problem structure presented in Glasserman and Tayur [4], which presents a single-product, multi-period, serial capacitated supply chain with production and inventory holding locations at each echelon. In their work, Glasserman and Tayur [4] use infinitesimal perturbation analysis (IPA) in order to determine optimal base stock levels in an order-up-to policy by optimizing over a sample path of the system.

Other approaches for solving the IMP have been reported in the literature. Chu et al. [5] use agent-based simulation-optimization on a multi-echelon system with an (r, Q) inventory policy. Expectations are determined via Monte Carlo simulation. Improvements are only accepted after passing a hypothesis test to mitigate the effect of noise on the improvement. Two-stage stochastic programming (2SSP) is used to optimize small supply chains in the works by Dillon et al. [6], Fattahi et al. [7], and Pauls-Worm et al. [8]. The studied supply chains are either single or two-echelon chains with centralized or decentralized configurations, a single perishable or unperishable product, and (r, S) or (s, S) policies. Zahiri et al. [9] present a multi-stage stochastic program (MSSP) for a four-level blood supply network with uncertain donation and demand. The model is reformulated and solved while using a hybrid multi-objective meta-heuristic. Bertsimas and Thiele [10] apply robust optimization to both uncapacitated and capacitated IMP. However, production capacity is not explicitly included. Their models are solved with linear programming (LP) or mixed-integer linear programming (MILP), depending on the usage of fixed costs. The reader is referred to Govindan and Cheng [11] for a review of robust optimization and stochastic programming approaches to supply chain planning.

Additionally, there have been a number of efforts to optimize multi-echelon supply chain problems via dynamic programming (DP). A neuro-dynamic programming approach was developed by Roy et al. [12] in order to solve a two-stage inventory optimization problem under demand uncertainty to reduce costs by 10% over the benchmarked heuristics. Kleywegt et al. [13] formulate a vendor managed inventory routing problem as a Markov Decision Process (MDP) and develop an approximate dynamic programming (ADP) method to solve it. Topaloglu and Kunnumkal [14] develop a Lagrangian relaxation-based ADP to a single-product, multi-site system to manage inventory for the network that outperforms a linear programming method used in the benchmark. Kunnumkal and Topaloglu [15] use ADP to develop stochastic approximation methods to compute optimal base-stock levels across three varieties of inventory management problems: a multi-period news vendor problem with backlogs and lost sales, and an inventory purchasing problem with uncertain pricing. Cimen and Kirkbride [16] apply ADP to a multi-factory and multi-product inventory management problem with process flexibility. They find that, in most scenarios, the ADP approach finds a policy within 1% of the optimal DP solution in approximately 25% of the time. Additional resources on supply chain management with DP and ADP is provided by Sarimveis et al. [17].

Reinforcement learning has also been applied to IMPs in recent years. Mortazavi et al. [18] use Q-learning for a four-echelon IMP with a 12 week cycle and non-stationary demand. Oroojlooyjadid et al. [19] train a Deep Q-Network in order to play the Beer Game—a classic example of a multi-echelon IMP—and achieve near optimal results. Kara and Dogan [20] use Q-learning and SARSA to learn stock-based replenishment policies for an IMP with perishable goods. Sultana et al. [21] use a hierarchical RL model to learn re-order policies for a two-level multi-product IMP with a warehouse and three retailers.

We extend the problem in Glasserman and Tayur [4] to general supply networks with tree topologies. Our focus is not on finding optimal parameters for static inventory policies, but rather to determine and compare different dynamic policy approaches to the IMP. We build on the previous works in the literature by exploring the IMP while using different approaches and discuss their relative merits and drawbacks. The approaches studied include

1. A deterministic linear programming model (DLP) that uses either the rolling horizon or shrinking horizon technique in order to determine optimal re-order quantities for each time period at each node in the supply network. Customer demand is modeled at its expectation value throughout the rolling/shrinking horizon time window.
2. A multi-stage stochastic program (MSSP) with a simplified scenario tree, as described in Section 2.7. Shrinking and rolling horizon for the MSSP model are both implemented to decide the reorder quantity at each time period.

3. A reinforcement learning model (RL) that makes re-order decisions based on the current state of the entire network.

We build off of the work of Hubbs et al. [22] by extending the IMPs presented therein in order to address multi-echelon problems with multiple suppliers at each echelon, and contribute new environments to the open-source OR-Gym project (See https://www.github.com/hubbs5/or-gym). The intial version of the IMP in the OR-Gym project was limited to serial multi-echelon systems and it did not include multi-stage stochastic programming models for reorder policy optimization. The library was thus generalized in order to simulate and optimize supply networks with tree topologies under uncertain demand, while using the dynamic reorder policies mentioned above.

2. Materials and Methods

2.1. Problem Statement

In this work, we focus on the multi-echelon, multi-period, single-product, and single-market inventory management problem (IMP) in a make-to-order supply network with uncertain stationary demand. The base case supply network has a tree topology with four echelons, as shown in Figure 1. The different sets that are used for the nodes in the base case network are designated in the figure's legend (raw material, J^{raw}; main, J; retail, J^{retail}; distributor, J^{dist}; producer, J^{prod}; and, market nodes, J^{market}).

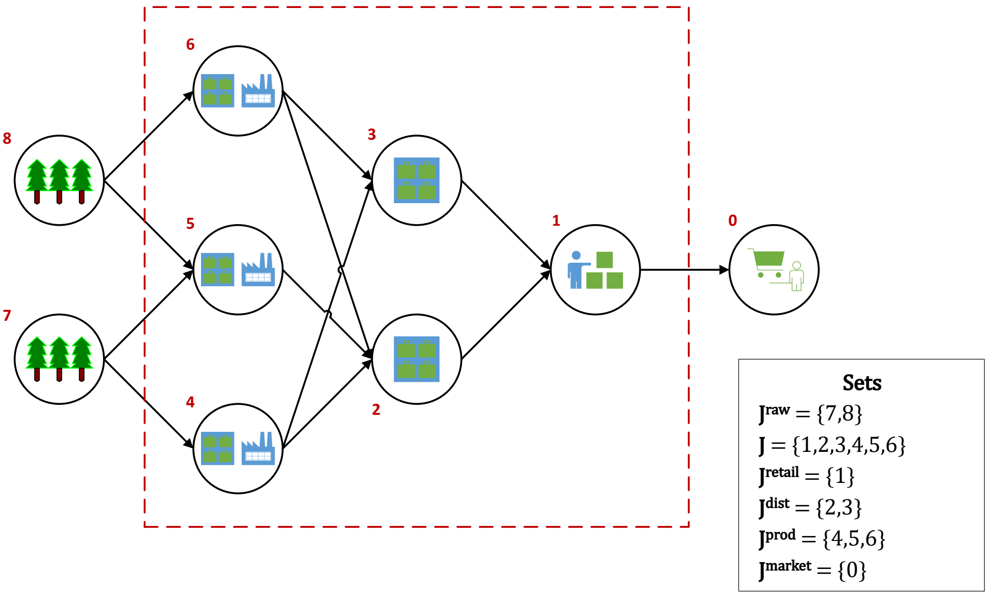

Figure 1. Supply Chain Network Schematic.

2.2. Sequence of Events

The sequence of events in each period of the IMP simulation environment occurs, as follows,

1. Main network nodes (retailer, distributors, and producers) place replenishment orders to their respective suppliers. Replenishment orders are filled according to available production capacity and available feedstock inventory at the respective suppliers. The supply network is assumed to be centralized, such that replenishment orders never exceed what can be provided by the suppliers to each node.
2. The main network nodes receive incoming feedstock inventory replenishment shipments that have made it down the product pipeline (after the associated lead times have transpired). The lead times between stages include both production times and transportation times.
3. Single-product customer demand occurs at the retail node and it is filled according to the available inventory at that stage.