Improving Inventory Management Through Simulation and Optimization

by

Yuwen Wang

and

Joshua Krause

Submitted on May 8th, 2026, in Partial Fulfillment of the
Requirements for the Degree of Master of Applied Science in Supply Chain Management
at the Massachusetts Institute of Technology

ABSTRACT

Newell Brands faces high working capital intensity due to a diverse consumer product portfolio and seasonal demand variability that creates high levels of on-hand inventory. To control costs, the company needs to optimize cycle stock parameters, specifically lot size and ordering frequency for raw materials, both of which are currently set by localized planner intuition. This project applies a simulation-based optimization framework to historical demand data for raw materials, using Monte Carlo methods to test various lot-sizing policies. The optimization algorithm identifies specific lot sizes that reduce total annualized inventory costs while maintaining required service levels, outperforming existing manual policies. We recommend that the supply planning organization adopt this as a decision support tool for regular inventory reviews to standardize raw material procurement decisions and leverage the identified cost trade-offs in future supplier negotiations.

Capstone Advisor: Dr. Selene Silvestri
Title: Research Scientist, MIT Center for Transportation & Logistics

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Background

Newell Brands is a global conglomerate that maintains a diversified portfolio of iconic consumer brands such as Marmot, Sharpie, Rubbermaid, Coleman, and Expo (Newell Brands, 2026). With multi-billion-dollar annual revenue (Newell Brands, 2026), Newell is constantly evaluating the efficiency of inventory management policies across its portfolio to improve financial performance.

1.2 Motivation

With a broad product portfolio containing a wide range of demand profiles comes an increase in supply chain complexity and working capital intensity driven by required service levels. Newell has faced challenges in forecasting accuracies for seasonal products and has been relying on safety stock to compensate for variability and to maintain service levels.

Historically, Newell has focused on safety stock optimization to drive efficiencies in its operating inventory. For this project, it seeks to leverage simulation and optimization engines to enhance its inventory control framework. This capstone focuses on analyzing cycle stock for a specific product line, examining ordering lot size and the resulting reorder point. The goal is to improve the company's inventory positions, reduce working capital, and gain insights about existing tradeoffs and bottlenecks in current supply planning parameters.

1.3 Problem Statement and Key Questions

To evaluate the optimal lot size and ordering frequency under Newell's current system, our project first reproduces the stochastic operational environment in a simulation engine and then proceeds to experiment with inventory policy optimization at scale.

Some key questions guiding our analysis include:

1. What supply planning parameters should be considered to optimize cycle stock for selected Stock Keeping Units (SKUs)?
2. How can stochastic simulation improve the robustness of supply planning and inventory policies?
3. What are key considerations for successful implementation of cycle stock controls in a complex network of interdependent manufacturing and distribution centers?

1.4 Project Goals & Expected Outcomes

By addressing these key questions, Newell Brands will be able to standardize the raw material ordering process across various production plants. As a result, procurement decisions will no longer be localized and executed solely based on the intuition of individual planners.
In the next section of our document, we cover several foundational aspects of inventory policy as well as key findings from literature that inform our project methodology.

2. STATE OF THE PRACTICE

To establish the methodology for our project, we begin by exploring existing literature to identify common challenges in inventory management and core levers of inventory policy design. Next, we review approaches for stochastic simulation and optimization principles. Finally, we survey techniques for improving the efficacy of inventory policy recommendations.

2.1 Common Challenges in Inventory Management

Our preliminary discussions with Newell stakeholders revealed that managing the flow of inventory involves a complex hand-off between demand forecasting, production planning, supply planning, and procurement. Through our literature analysis, we identified three prevalent challenges encountered when developing new inventory management solutions.

Errors and uncertainty in demand forecasting. Managing inventory to support sales requires granular knowledge of downstream customer demand. Despite advanced statistical techniques and primary data sources that can approximate future demand, the inherently stochastic nature of demand leads to over- or under-estimation — a challenge highlighted in a recent systematic review of uncertainty in supply chain optimization (Chen et al., 2023). This can be further exacerbated by highly seasonal products, where consumer demand can spike or dip suddenly. Taken together, these uncertainties create difficulty in determining the appropriate inventory policy for a given product.

Multi-echelon network complexity. In any multi-tier network, an individual node is dependent on the performance of other nodes, and fluctuations can have cascading impacts up- and downstream (Taskin, 2009). Additionally, orders at each node of the supply chain may be batched together, creating lumpy demand signals for upstream nodes and leading to excess inventory build-up.

Working capital optimization. Companies must make careful tradeoffs between investing capital into inventory production – and subsequently risking obsolete inventory – vs. facing a stock-out and missing out on potential customer revenue. Even when inventory is non-perishable, it can become obsolete if customer preferences shift or manufacturing processes change. Careful inventory management must ensure that working capital is deployed in a manner that supports the financial health of the company and aligns with strategic goals (Bendavid et al., 2017).

2.2 Inventory Policy Design: Safety Stock and Cycle Stock

Companies can employ several measures to protect against the inherent variability of real-world operations and address the challenges discussed in Section 2.1. The most common technique is to hold safety stock (Muller, 2011). This serves as a buffer to protect against unexpected fluctuations in demand and ensures a company meets its established service level.

While safety stock directly addresses uncertainty, cycle stock represents planned inventory for expected demand. Cycle stock is the portion of inventory held to meet expected demand between replenishment cycles (Chopra & Meindl, 2013), and it is primarily determined by two factors: lot size and order frequency.

Once safety stock has been optimized, managing cycle stock is the main lever for balancing cost efficiency and service-level performance.

2.3 Managing Cycle Stock

Lot size is the most direct driver of cycle stock. In this project, ordering cost S is modeled as a per-lot charge: S is incurred for each lot of Q units dispatched in a replenishment event, so an event covering a deficit of n lots incurs $n \times S$. Larger orders therefore reduce total ordering cost per replenishment cycle but increase holding costs. Smaller, more frequent orders reduce average inventory and holding costs but increase total ordering costs (Chopra & Meindl, 2013). This trade-off is further complicated by supplier-imposed minimum order quantities (MOQs), which set a lower bound on order size and can force companies to hold more inventory than desired.

Order frequency determines how often replenishment occurs and interacts closely with lot size to shape cycle stock levels. Two common approaches govern this process: continuous review and periodic review. In continuous review systems, inventory is monitored in real time, and orders are triggered when stock reaches a predefined reorder point (the inventory level that triggers a replenishment order). In periodic review systems, orders are placed at fixed intervals, with order quantities adjusted based on the actual demand observed during the review period. Each method carries distinct implications for cost structure, responsiveness, and inventory stability (Chopra & Meindl, 2013).

2.4 Simulation-Based Optimization for Improved Inventory Management

Building on this foundation of inventory policy design, we then analyze existing literature to validate Newell's desire to leverage simulation-based optimization and generate recommendations for their chosen product line.

Analytical models for supply chains often use system dynamics to illustrate the operational impacts of up- and down-stream tiers at a given node. System dynamics-based approaches are most suitable for capturing interactions through feedback loops and evaluating long-term behavior, while simulation (e.g., Monte Carlo) helps to identify improvements to baseline policy (Law & Kelton, 2007). Because our primary goal for Newell is to evaluate the performance distribution of alternative lot-sizing options under uncertain conditions, a simulation-based optimization approach is more appropriate and analytically defensible (Law & Kelton, 2007).

A prerequisite for constructing a valid simulation environment is to identify the probability distribution that describes historical data. To tackle this, researchers must employ a combination of summary statistics and goodness-of-fit testing to identify the best-fit probability distribution(s) for each variable (Law & Kelton, 2007). The distribution parameters identified in the result are then used to generate synthetic data that will mirror the stochastic nature of the historical data. Leveraging the full empirical dataset for distribution fitting ensures the underlying stochastic behavior is accurately captured and reduces bias introduced by finite sample sizes (Hastie et al., 2009).

Before synthetic demand can be generated, the historical demand series must be assessed for stationarity: the property of having a stable mean and variance over time. Applying a distribution fit directly to a non-stationary series conflates long-run structural patterns, such as trend and seasonal cycles, with the residual week-to-week variability that the simulation should model. The Augmented Dickey-Fuller (ADF) test provides a formal diagnostic for this condition: when it signals non-stationarity, the series must be decomposed before any distribution can be fitted to the noise component (Dickey & Fuller, 1979). Seasonal-Trend Decomposition using LOESS (STL) addresses this by separating the demand series into three additive components: a smooth trend, a repeating seasonal cycle, and a stochastic residual (Cleveland et al., 1990). This ensures only the unexplained variation is passed to the distribution-fitting step.

Once isolated, the residuals must satisfy a further condition: they should be independently and identically distributed (i.i.d.). The Ljung-Box test diagnoses whether residuals exhibit serial autocorrelation — systematic dependence between consecutive observations — which, if present, would invalidate independent sampling from a fitted distribution (Ljung & Box, 1978). When the test confirms independence, the best-fit parametric distribution can be selected by ranking candidates using the Akaike Information Criterion (AIC), which penalizes model complexity while rewarding goodness of fit (Hyndman & Athanasopoulos, 2021). When autocorrelation is detected, a non-parametric moving block bootstrap can be used instead, which resamples consecutive blocks of empirical residuals to preserve their temporal structure (Künsch, 1989).

Taken together, these findings validated our decision to employ a decomposition-based simulation framework and directly inform the methodological design described in Section 3.

3. METHODOLOGY

As discussed in Section 2.4, our literature review points to a multi-step simulation and optimization process for yielding recommendations on optimal inventory policy. Table 1 illustrates the sequence of our project, including the key inputs and outputs:

Steps	Input	Process	Output	Purpose
1. Data Collection and Preparation	Empirical data *(see Table 2)*	Clean and normalize data	Processed data for analysis	Eliminate errors in subsequent analysis
2. Distribution Fitting	Processed data	Assess stationarity via ADF test; decompose via STL. Validate residuals using ADF (stationarity) and Ljung-Box (independence) to determine simulation path.	Simulation path selection: parametric distribution fit (AIC-ranked via Fitter) if residuals are i.i.d.; moving block bootstrap if serial autocorrelation is detected.	Establish the stochastic mechanism for synthetic demand generation based on observed residual properties.
3. Simulation	Corresponding demand characteristics (Trend, Seasonality, Residual) to process STL Re-composition.	Use best-fit method for the residuals to simulate theoretical demand values by STL re-composition.	Simulated weekly demand series.	Generate stochastic demand realizations to quantify policy performance under uncertainty
4. Optimization	Simulated demand data, inventory policy parameters	Perform optimization and evaluate corresponding total relevant cost for each scenario	Optimal lot size (constrained vs. unconstrained)	Identify optimal lot size based on total relevant cost distribution, subject to service level constraint
5. Stress Testing	Optimal lot size results	Stress test each inventory policy to evaluate the long-term effect on total cost.	Evaluate the total cost difference between current, unconstrained, and constrained solutions	Ensure the total cost reflects the true operational impact

Table 1. Project Approach

3.1 Data Collection and Preparation

Upon receiving the datasets shown in Table 2, preliminary cleaning is necessary to mitigate potential downstream errors while generating synthetic data. We conduct a systematic assessment of data quality issues, including the presence of null or missing values, redundant records, formatting inconsistencies, and extreme outliers.

Entity	Attributes	Time Horizon	Purpose in Project
Material Master	Product SKU, Plant, Lot Size, Storage Constraints, Lead Time Variability, Order Cost Parameter	Current (snapshot)	Defines core planning parameters used in the total cost and cycle stock optimization model; provides static reference data for all SKUs
Inventory Data	Period Inventory Levels, Safety Stock, Cycle Stock vs. Safety Stock Split, Inventory Value ($)	12 months back	Captures historical inventory dynamics to calibrate and validate the simulation model; forms the baseline for service level and cost analysis
Demand Data	Historical Demand, Order Actuals, Forecast Demand, Forecast Error or Bias (%), Promotional or Seasonal Flags	24 months back, 12 months forward	Provides stochastic input for demand uncertainty modeling; supports fitting probability distributions and simulation of demand variability
Production – Sales – Inventory (PSI) Data	Lead Time Variability, Service Level Targets (per SKU)	12 months back, 12 months forward	Enables validation of planned vs. actual performance; aligns policy recommendations with operational and service level targets
Cost Parameters	Holding Cost, Order Cost, Stockout Cost, Transportation Cost	Current (snapshot)	Provides economic levers for total cost optimization and simulation of trade-offs between holding, ordering, and service costs
Supporting / Other	Purchase Order History, Supplier Performance Metrics	TBD	As needed; Enhances model calibration by capturing real-world supply variability and supplier reliability factors; used in advanced simulations

Table 2. Primary Data from Newell

For the purposes of this project, we had access to two years of historical demand data and one year of forecasted demand, comprising 155 weeks of weekly demand for the pilot material.

3.2 Simulation Method Selection

Before synthetic demand can be generated, a method must be chosen for representing the stochastic behavior of the demand series. This section describes the candidate approaches considered and explains why STL decomposition was selected as the most appropriate method given the characteristics of Newell's data. The choice directly shapes the quality and reliability of every simulation output that follows.

The central challenge is that real-world demand contains two fundamentally different types of variation. The first is deterministic structure: the predictable trend and seasonal patterns that repeat year over year. The second is stochastic noise: the week-to-week randomness that a simulation should actually model. Failing to separate the residual noise from the predictable trend produces synthetic demand that overstates uncertainty. For example, fitting a distribution directly to the full range of raw weekly demand values treats seasonal highs and lows as if they were random fluctuations, inflating the noise estimate and producing unrealistically volatile simulated demand (Hyndman & Athanasopoulos, 2021). A bootstrap applied to the raw series has the same problem: the resampled blocks carry seasonal structure along with the noise, so the synthetic series does not faithfully represent the uncertainty a planner actually faces week to week. Both simple approaches were evaluated and ruled out on this basis.

Decomposition-based methods address this by separating the demand series before simulation: Trend and seasonal patterns are captured in a deterministic schedule and only the remaining residual (the unexplained variation) is treated as stochastic. Four decomposition-based approaches were evaluated: a rolling mean smoother with parametric residual fitting, a rolling mean smoother with ARMA residual modeling, Fourier regression with parametric fitting, and Fourier regression with ARMA modeling. Each captures the deterministic structure differently and makes different assumptions about the shape of the trend and seasonal cycle. ARMA residual modeling was tested because it can improve confidence intervals when residuals exhibit serial autocorrelation, though this advantage disappears when residuals are effectively random noise and the tighter intervals reflect absorbed autocorrelation rather than genuine stochastic reduction.

STL (Seasonal-Trend Decomposition using LOESS) was selected as the most suitable approach for Newell's demand data (Cleveland et al., 1990). The key distinction is flexibility: unlike rolling mean or Fourier methods, STL makes no assumption about the shape of the trend or the harmonic structure of the seasonal cycle, allowing each to evolve naturally from the data. Its robust fitting variant further down-weights outlier weeks, preventing atypical demand spikes from distorting the estimated components. Most importantly, STL's accuracy improves with the number of complete annual cycles observed. With three years of weekly data available for this project, the seasonal component is estimated from three independent realizations of the annual pattern, producing a more stable and generalizable decomposition than approaches that require only a single cycle and directly leveraging the breadth of available historical data.

Once STL residuals are separated, a two-step diagnostic determines how they are simulated. The ADF test first confirms that the residuals are stationary. The Ljung-Box test then evaluates whether they are serially independent. If they are, the best-fit parametric distribution is selected by AIC using the Fitter library. If serial autocorrelation is present, as was the case with the pilot material, a moving block bootstrap method is applied instead, which resamples blocks of consecutive empirical residuals to preserve their temporal structure (Künsch, 1989). This routing logic ensures the simulation adapts automatically to the stochastic properties of each material without requiring manual configuration.

STL with conditional routing between parametric and non-parametric paths was selected as the final demand generation method, based on its methodological rigor, robustness to the residual properties observed in the pilot data, and the interpretability of its components. The trend, seasonality, and residual outputs are directly reportable as diagnostics, providing transparency into what the model learned from the data before any simulation is run.

3.3 Distribution Fitting

Following the approach justified in Section 3.2, we first apply the ADF stationarity test to the historical weekly demand series as a routing diagnostic. We decompose demand using STL with a 52-week period and robust outlier handling. With only three available annual cycles, permitting the seasonal component to evolve linearly across years would over-fit the boundary years and dump unexplained variance into the middle of the series. Constraining the seasonal pattern to be stable across years yields a residual series whose variance is distributed evenly across the full historical window.

To maintain robustness, the modeling logic for distribution fitting incorporates both parametric and non-parametric paths, with a multi-step diagnostic procedure to determine the appropriate path for each material, as outlined in Image 1.

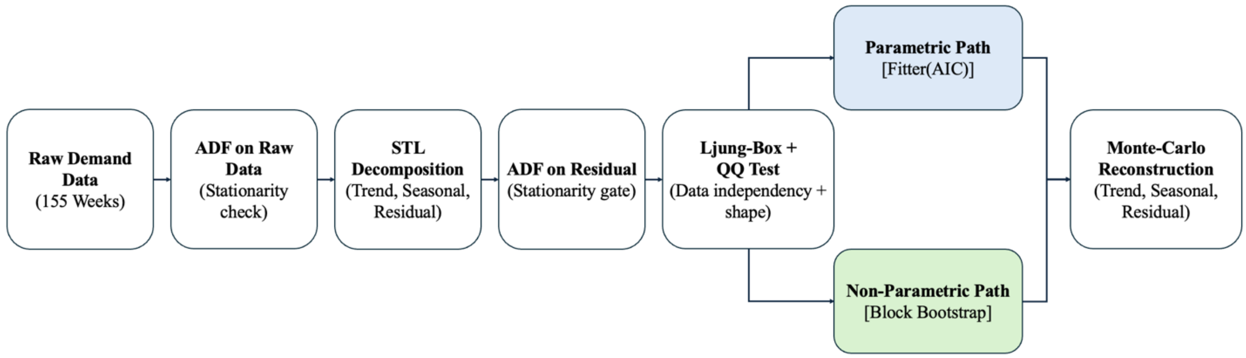

Image 1. Distribution Fitting Methodology

First, the ADF test confirms stationarity of the residuals (a necessary precondition for any distribution-fitting). Then, the Ljung-Box test at lags 5, 10, and 20 evaluates whether residuals are independent across time. Finally, as a diagnostic step the QQ plot (Quantile-Quantile Plot) provides a visual check on tail behavior, which matters disproportionately for inventory because stockout risk is driven by the right tail of the demand distribution.

Structurally, the demand generator is constructed along two parallel paths, with the diagnostic outcomes determining the appropriate path. Both paths share an identical deterministic schedule, built from the STL components: the trend is extrapolated forward 52 weeks using a linear fit through the most recent year of the trend series, and the seasonal cycle is repeated from the most recent observed annual pattern. This schedule represents the expected demand level for each future week.

For the pilot material, the active path is the block bootstrap. Image 2 shows the residual diagnostic suite, with the time series, autocorrelation function, rolling standard deviation, and QQ plot against the best-fit parametric reference.

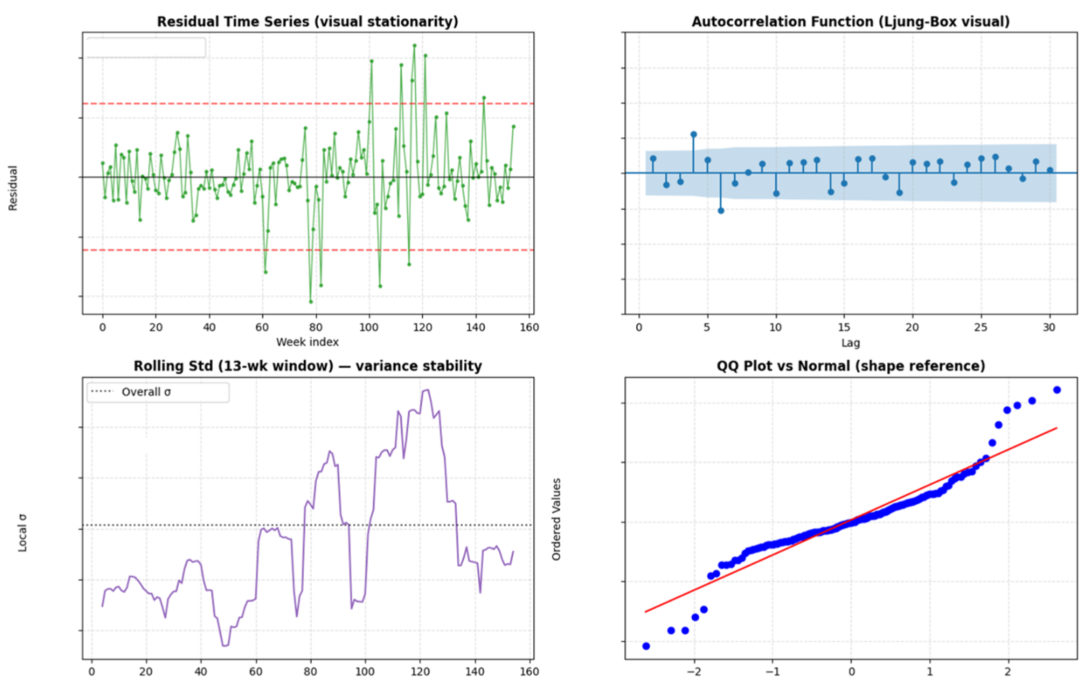

Image 2. Ljung-Box and QQ Plot Analysis on Residual.

Simulated weekly demand for each Monte Carlo iteration is created through reconstruction of trend, seasonality, and residual, where the residual term is determined based upon whichever path the diagnostics selected. The parametric machinery is retained even when the bootstrap path is active, with the AIC ranking and QQ plots reported as diagnostic context; they document how the parametric path would have performed and provide a record of the tail mismatch that motivated the routing decision.

3.4 Formal Model Specification

This section introduces compact notation to define the simulation model. All symbols used in subsequent sections are grounded here.

State Variable

Let IP_t denote the inventory position at the end of week t. It is defined as on-hand inventory after demand is consumed plus the total quantity outstanding in all in-transit replenishment orders:

$$IP_t = OH_t + \sum_{j \in P_t} q_j$$

where OH_t is the physical on-hand inventory after demand for week t is fulfilled, P_t is the set of pending orders, and q_j is the quantity of order j. When OH_t falls below zero, the negative quantity is recorded as an unfilled unit (shortage); for holding-cost computation, on-hand inventory is floored at zero.

Reorder Point

The reorder point s is defined using the standard sigma-based formulation:

$$s = \mu^L + z_\alpha \cdot \sigma^L$$

where $\mu^L = L \cdot \mu_u$ is expected demand over the replenishment lead time of L weeks, $\sigma^L = \sigma_u \cdot \sqrt{L}$ is the standard deviation of lead-time demand under the i.i.d. weekly demand assumption, μ_u and σ_u are the mean and standard deviation of the fitted weekly demand distribution, and z_α is the safety factor corresponding to service level α. In the implementation, the safety buffer is parameterized as $k \cdot \mu_u$ where k denotes safety stock coverage in weeks, which is equivalent to the sigma-based formulation when $z_\alpha \cdot \sigma_u = k \cdot \mu_u$.

Ordering Policy

The model implements a continuous-review (s, nQ) policy as described in Section 2.3, where s is the reorder point, Q is the fixed lot size, and n represents the number of lots ordered. After demand is consumed in each week t and scheduled receipts are credited, the inventory position is evaluated. If $IP_t < s$, a replenishment order is triggered for n_t lots of size Q, restoring IP_t to at least s:

$$n_t = \left\lceil \frac{s - IP_t}{Q} \right\rceil$$

The order arrives L weeks later. Each lot incurs the fixed ordering cost S; a single replenishment event ordering n_t lots therefore incurs $n_t \cdot S$. This per-lot cost structure reflects Newell's operational context, in which each discrete lot represents an independent shipment unit carrying a fixed cost.

Cost Function

The optimizer minimizes expected Total Relevant Cost (TRC) over the 52-week horizon, comprising holding and ordering costs only. Purchasing cost C is excluded because it is proportional to total demand volume and independent of lot size Q:

$$\mathbb{E}[TRC(Q)] = H \cdot \mathbb{E}[\bar{I}_t] + S \cdot \mathbb{E}[N_t]$$

where:

$$\bar{I}_t = \frac{1}{T}\sum_{t=1}^{T} \max(0, OH_t)$$

$$N_t = \sum_{t=1}^{T} n_t$$

In these formulas, $\bar{I}_t$ is average on-hand inventory over $T = 52$ weeks, N_t is total lots ordered, H is the annual holding cost rate in dollars per unit per year, and S is the per-lot ordering cost in dollars per lot. Average on-hand inventory is computed directly from the simulated weekly inventory trace rather than approximated as $Q/2$.

Optimization Objective

The formal problem is to minimize the expected total relevant cost:

$$\min_{Q} \mathbb{E}[TRC(Q)]$$

subject to the service level constraint:

$$\text{IFR}(Q) \geq \tau$$

where the Item Fill Rate (IFR) is defined as:

$$\text{IFR}(Q) = 1 - \frac{\mathbb{E}[U]}{\mathbb{E}[D]}$$

Here, U is total units short over the horizon, D is total units demanded over the horizon, and τ is the user-specified service level target. Candidate lot sizes for which the mean simulated IFR falls below τ are assigned infinite cost and excluded from selection. The search algorithm and convergence criteria are described in Section 3.5.

All fixed parameters for the pilot analysis are listed in Table 3. Parameters labeled "Newell" are sourced directly from Newell's material master and cost accounting systems; parameters labeled "Analyst" reflect modeling choices calibrated to the available data and validated through the simulation framework. Parameters labeled “Reproducibility” are fixed solely to ensure identical results across simulation runs.

Parameter	Symbol	Basis
Ordering cost (per lot)	S	Newell
Annual holding cost rate	H	Newell
Unit purchasing cost	C	Newell
Replenishment lead time (deterministic)	L	Newell
Safety stock coverage	k	Analyst
IFR service level target	τ	Newell
Monte Carlo iterations per candidate Q	N	Analyst
Simulation horizon	T	Analyst
Initial on-hand inventory	I_0	Newell
Random seed	—	Reproducibility

Table 3. Model Parameters and Assumptions

3.5 Simulation and Optimization

Building on the insights gained through distribution fitting, our model then generates synthetic data representing possible 52-week demand periods. After these plausible demand distributions are created, the simulation and optimization procedures are executed in sequence. The synthetic demand samples are propagated through repeated simulation runs, during which the performance of the existing inventory policy is evaluated across a range of lot-size choices. For each lot size, the system conducts 1,000 simulation replications, yielding a distribution of performance outcomes specific to that lot-size configuration.

The optimization employs a multi-phase adaptive search algorithm. In each phase, 21 candidate lot sizes are evaluated across a window centered on the current best estimate. The algorithm tests 10 increments below and 10 above the center point, plots the resulting cost curve, and evaluates whether the minimum falls within a three-position buffer of the center. If the minimum is off-center, the window shifts and re-evaluates; if centered, the algorithm advances to the next phase with a reduced step size (80% of the prior step). This process continues until the step size converges to within 1% of the initial lot size, producing a fine-grained optimum. Boundary conditions are enforced to ensure lot sizes remain positive and do not exceed the historical maximum weekly demand.

By iterating this process across all feasible lot-size alternatives, we obtain distributions of expected total cost, enabling comparative evaluation and selection of the optimal lot-size policy. Then based on the average performance across the 1,000 iterations, the user input service level is used as a filter to eliminate any candidate lot sizes that do not meet the minimum service level requirement.

The entire process allows end-users to experiment with different ordering parameters such as ordering cost, holding cost, and service level to gain direct, empirical insight into inventory trade-offs without having to experiment in a real-world environment. As a benchmark, the classical Economic Order Quantity (EOQ) is also computed to validate the simulation's directional output.

4. RESULTS AND DISCUSSION

This section presents the results of the simulation-based optimization framework applied to Newell's chosen product portfolio. We first describe the outputs of the distribution fitting and Monte Carlo simulation process, then discuss the business implications for Newell's inventory planning organization and finally acknowledge key limitations of the current analysis.

4.1 Results

The analysis proceeds through five sequential stages as outlined in our methodology: data preparation, distribution fitting, Monte Carlo simulation, optimization, and comparative stress testing.

Data Preparation and Distribution Fitting

Following the data preparation steps described in Section 3.1, the pilot yielded a cleaned 155-week weekly demand series spanning 2024–2026. The model identifies the corresponding trend, seasonality, and residual as illustrated in Image 3.

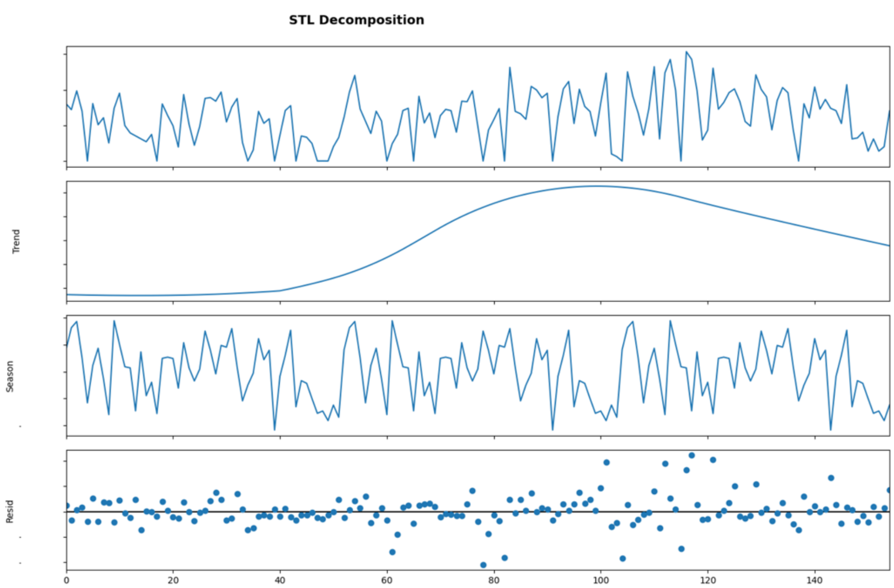

Image 3. STL Decomposition Analysis on Trend, Seasonality, and Residual.

Following the decomposition of the weekly demand series into trend, seasonality, and residuals, the residuals undergo a diagnostic assessment to dictate the subsequent simulation trajectory. For the pilot material, the ADF test on the raw 155-week series returns a p-value above the 0.05 threshold, confirming non-stationarity. This result is consistent with visual inspection: the demand exhibits a clear upward trend across 2024–2026 and a recurring annual seasonal pattern, both of which violate the constant-mean assumption underlying any direct distribution fit. Decomposition is therefore required before any noise model can be calibrated. This stochastic generation of residuals facilitates the final re-composition of synthetic demand data, shown in Image 4, ensuring that the simulated output maintains the structural integrity of the original time-series components.

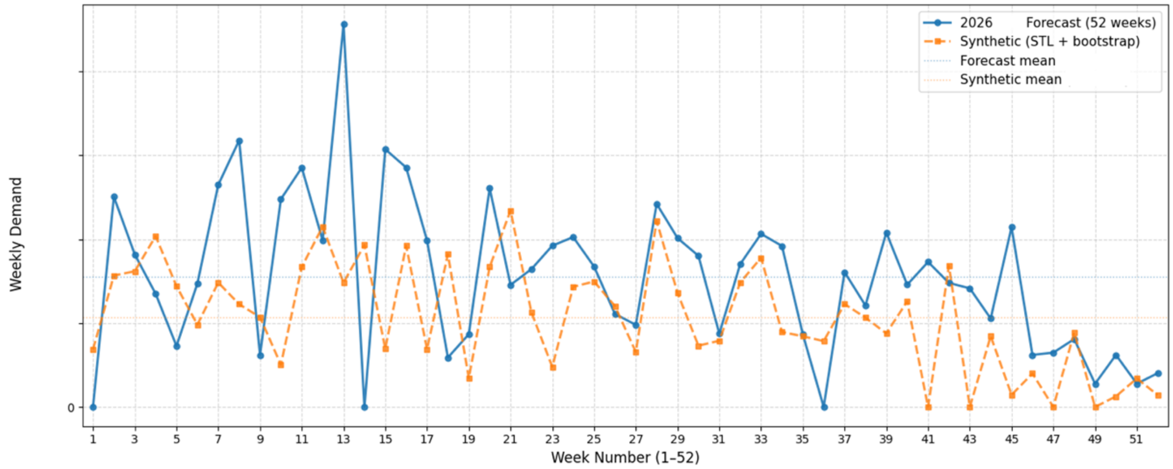

Image 4. 52-Week Forecast vs. Synthetic Demand

Simulation and Lot-Size Optimization

The core simulation engine evaluates candidate lot sizes over a 52-week time horizon using 1,000 Monte Carlo iterations per candidate. Each iteration generates synthetic weekly demand, simulates inventory dynamics, and computes annualized total relevant cost—comprising ordering costs and holding costs. Ordering logic incorporates a deterministic lead time. Average inventory is computed directly from the simulated weekly inventory history, which reflects actual system behavior rather than the classical Economic Order Quantity (EOQ) approximation of lot size / 2.

The resulting cost curve, shown in Image 5, exhibits a global minimum total relevant cost, with a characteristic U-shape predicted by inventory theory: small lot sizes incur high ordering costs, while excessively large lot sizes drive holding costs up (Chopra & Meindl, 2013). The algorithm consistently identifies an interior optimum that balances these competing cost components. The optimization converged to an optimal lot size that was over 5x smaller than the planner specified lot size.

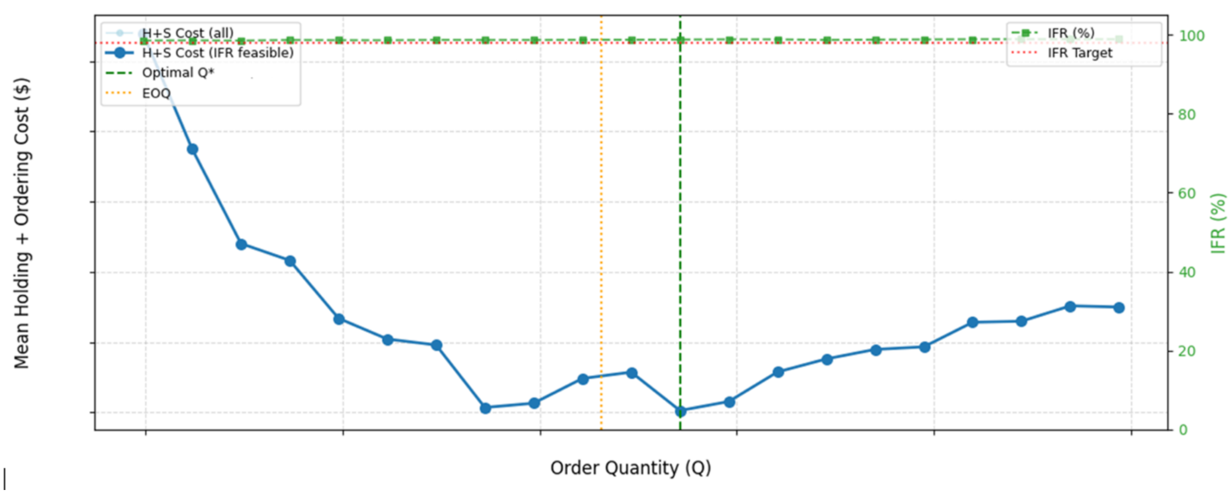

Image 5. Constrained Optimization and Total Relevant Cost (H+S)

Stress Testing and Validation

To validate robustness, a head-to-head stress test compares the initial lot size against the algorithm-optimized lot size over 1,000 Monte Carlo iterations. Both policies are subjected to identical demand scenarios using a fixed random seed, ensuring a controlled comparison. The stress test records the mean annualized total relevant cost and stockout frequency for each policy. Results confirm that the optimized lot size yields a lower mean total cost relative to the baseline lot size while maintaining an acceptable service level. The cost comparison output quantifies total relevant cost (annual ordering + holding) savings of 14.6% against baseline, providing a clear decision metric for the inventory planning team.

Image 6 provides visual confirmation of the reduced ordering and holding cost when smaller lot sizes keep the total inventory position (on-hand inventory + in-transit inventory) much closer to the defined reorder point.

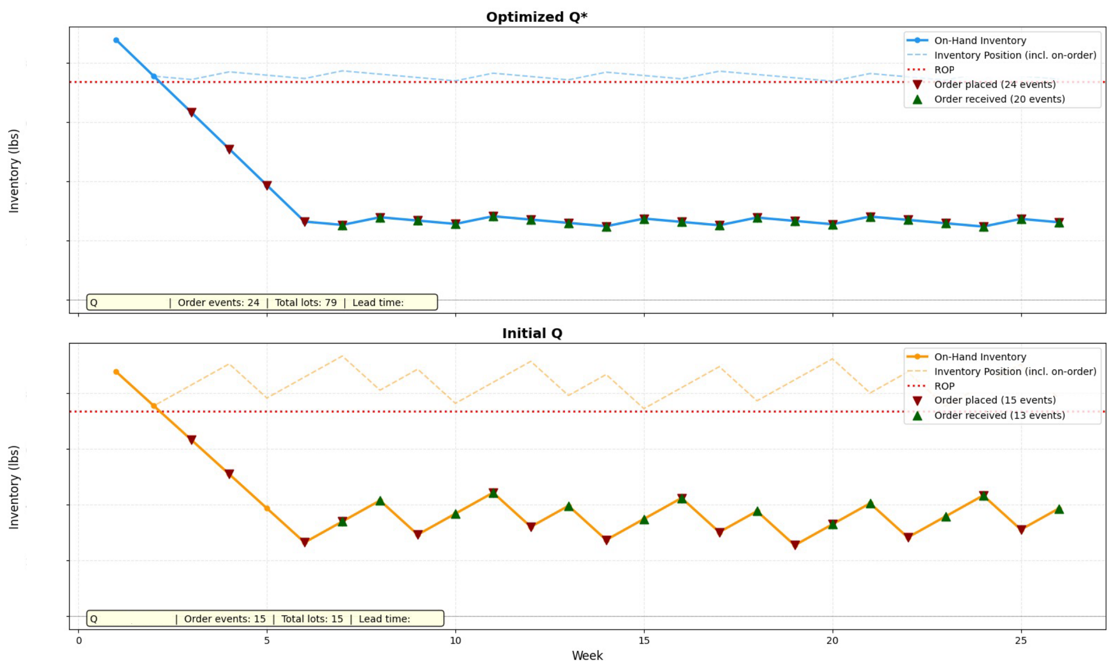

Image 6. Optimized Inventory Performance vs. Baseline

EOQ Benchmark Comparison

As a sanity check on the simulation output, the classical EOQ is computed using the formula:

$$Q^* = \sqrt{\frac{2DS}{H}}$$

where Q* is the optimal lot size, D is the estimated annual demand, S is the ordering cost, and H is the annual holding cost per unit. The EOQ represents the theoretically optimal lot size under deterministic demand with no service-level constraints.

While our exact parameter values for the pilot material have been removed for confidentiality, the Q* represented a 10% increase above the EOQ benchmark.

The Monte Carlo-optimized lot size diverges from this EOQ benchmark for two structurally meaningful reasons. First, the EOQ model assumes demand arrives at a constant, deterministic rate, while the simulation operates under stochastic weekly demand governed by the STL-fitted distribution. Under uncertainty, larger safety buffers and more conservative ordering behavior shift the cost-minimizing quantity upward. Second, the simulation enforces an explicit Item Fill Rate constraint: candidate lot sizes that fail to meet the IFR threshold are assigned infinite cost and excluded from selection, regardless of their holding and ordering cost profile. This service-level floor induces a preference for larger order quantities relative to the deterministic optimum. Taken together, these factors explain and justify the divergence between the EOQ benchmark and the simulation result, confirming that the simulation is correctly penalizing policies that sacrifice service level in pursuit of short-run cost minimization.

Service Level Sensitivity Analysis

Image 7 reports the IFR sensitivity sweep across targets from 90% to 99%.

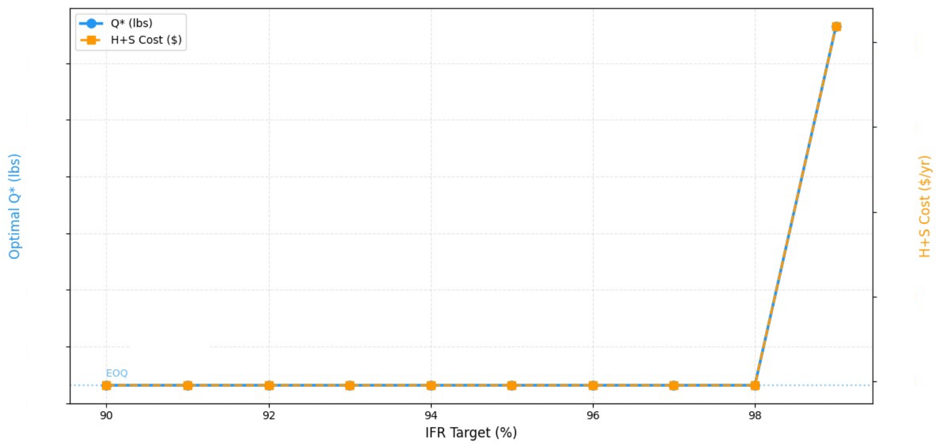

Image 7. Service Level Impact on Q* and Cost

A notable structural feature emerges: the cost-minimizing lot size is unchanged for every target up to 98%, because the unconstrained optimum already delivers a sufficient simulated fill rate — therefore the IFR constraint is non-binding across this range. Targeting 99% exacts a substantially larger penalty, with Q* expanding to 2.91× EOQ. The convex, hockey-stick shape of the Q*–IFR relationship reflects a fundamental trade-off in continuous review systems: high fill rates require large lot sizes to reduce the probability of stockout during the replenishment lead time, and the marginal cost of each incremental service point accelerates sharply in the upper tail.

4.2 Implications

The results of this analysis carry several significant implications for Newell's inventory planning organization and broader supply chain strategy.

Working Capital Reduction Through Cycle Stock Optimization

Currently the model identifies a 14.58% reduction in total relevant cost — comprising ordering and holding costs — for the pilot material. Of that amount, holding costs specifically are reduced by 17%. Because holding cost, not ordering cost, scales with inventory volume across the portfolio, the holding cost reduction is the more relevant figure for estimating enterprise-wide savings. Applying this rate to Newell's reported inventory balance at their benchmark annual holding cost rate implies significant holding cost savings opportunity if comparable optimization headroom exists across the entire portfolio. The realized savings will vary by material depending on how far current planner-specified lot sizes deviate from their simulation-optimized counterparts, a figure that will become clearer as the framework is extended to the full portfolio.

The simulation demonstrates that the existing planner-specified lot sizes are not necessarily aligned with the cost-minimizing optimum for the chosen product. By right-sizing order quantities through simulation-based optimization, Newell stands to reduce average on-hand inventory and the associated holding costs. Because a few raw materials can represent a substantial share of input spend for the chosen product, even modest percentage improvements in cycle stock efficiency translate into meaningful working capital release. This freed capital can be redeployed toward higher-priority investments or used to reduce short-term financing costs.

Standardization of Supply Planning Decisions

A key challenge identified during stakeholder interviews is that decisions on ordering amount are historically made at the discretion of individual planners, resulting in inconsistent policies across materials and plants. The simulation framework introduced in this project provides a data-driven, repeatable methodology for evaluating lot-size alternatives. By parameterizing the tool with material-specific cost inputs and service level, Newell's planning team gains the ability to generate defensible lot-size recommendations for any material in the portfolio. This standardization reduces reliance on planner intuition and creates a shared analytical foundation for procurement decisions.

Visibility Into Cost Trade-Offs and Binding Constraints

The U-shaped cost curves generated during the multi-phase search provides direct visibility into the trade-off between ordering frequency and inventory carrying costs. These visualizations illustrate why a given lot size is preferred and how sensitive the total cost is to deviations from the optimum. Furthermore, the model allows users to experiment with various parameters to conduct sensitivity analysis and gain empirical insights that enable prioritization of operational improvements and supplier negotiations with the greatest potential return.

Service Level Resilience

The stress test results demonstrate that the optimized lot sizes do not sacrifice service-level performance in pursuit of cost savings. Stockout frequency under the optimized policy is comparable to the initial baseline policy across simulated demand scenarios. This finding is particularly important for Newell leadership, as any recommended change to ordering parameters needs to preserve the company's commitment to demand fulfillment. The simulation framework thus provides assurance that cost optimization and service-level maintenance are not mutually exclusive objectives.

4.3 Limitations

While the simulation framework produces actionable results, several limitations should be acknowledged when interpreting the findings.

First, the lead time is modeled as deterministic. In practice, supplier lead times exhibit variability driven by production scheduling, transportation disruptions, and capacity constraints. Incorporating stochastic lead times would produce a more conservative and realistic assessment of safety stock requirements and stockout risk.

Second, the demand distribution relies on the accuracy of historical and forecasted demand data for generating feasible synthetic demand. Consequently, the framework may fail to account for unprecedented market disruptions or structural shifts that are not represented in the available dataset. This inherent dependency on past performance potentially limits the model's robustness when faced with 'black swan' events or evolving consumer behaviors that deviate from observed statistical norms.

Third, the simulation evaluates each material in isolation and does not account for interdependencies across Newell's multi-echelon network. Production schedules, shared warehouse capacity, and joint procurement agreements create coupling effects that may constrain the feasibility of implementing individually optimized lot sizes.

Finally, the current analysis is executed on a per-material basis with manual parameter entry via interactive widgets. While this approach enables rapid prototyping and validation, scaling the framework to the full material portfolio will require automation of the input pipeline and integration with Newell's enterprise data infrastructure.

5. RECOMMENDATIONS

Based on the results and implications discussed in Section 4, we offer the following recommendations for Newell's inventory planning leadership. These recommendations are organized into near-term management actions and longer-term opportunities for extending the analytical framework.

5.1 Management Recommendations

We recommend that Newell's supply planning team adopt the simulation-based lot-size optimization tool as a standard component of the continuous planning review process for the chosen product line. Specifically, we propose the following actions.

First, Newell should execute the simulation pipeline across the remaining materials in the product portfolio to identify lot-size optimization opportunities for each SKU. The model is designed with configurable input widgets for inventory and simulation parameter values, enabling planners to run the analysis without modifying the underlying code. As each material is evaluated, the planning team should document the optimized lot size, the magnitude of projected annual savings, and any binding constraints surfaced by the cost curve analysis.

Second, we recommend establishing a regular cadence for re-running the optimization as new demand data becomes available. Demand patterns for materials may shift due to promotional activity, product mix changes, or macroeconomic factors, and a periodic refresh ensures that lot-size recommendations remain aligned with current conditions. The simulation's relatively low computational overhead makes this a feasible addition to the existing planning workflow.

Third, we recommend dynamically evaluating penalty cost to determine the corresponding service level target at the raw material level. The result will yield a more rigorous approach in determining the right safety stock level corresponding to the established service level for each material. This allows the optimization framework to optimize overall inventory position in a holistic manner and identify additional cost saving opportunities.

Finally, where the sensitivity analysis reveals that supplier-imposed minimum order quantities or transportation capacity constraints are the binding factors preventing further cost reduction, we recommend that Newell engage in targeted supplier negotiations. Armed with the quantified cost impact of these constraints, as demonstrated by the simulation's sensitivity outputs, the procurement team would be positioned to negotiate more favorable terms or explore alternative logistics configurations.

5.2 Future Work

Several avenues for extending this work would enhance the framework's accuracy, scalability, and organizational impact.

The most immediate extension would be to incorporate stochastic transportation lead times into the simulation engine. Replacing the current deterministic replenishment lead-time assumption with a probability distribution fitted to historical supplier performance data would enable more realistic modeling of inventory position variability and produce more robust safety stock recommendations. This enhancement would be particularly valuable for materials sourced from suppliers with variable delivery performance.

A second direction involves extending the framework to model multi-echelon interdependencies. The current tool optimizes lot sizes for individual materials in isolation, but in practice, raw material ordering decisions affect downstream production scheduling, finished goods inventory at distribution centers, and ultimately, retailer service levels. A multi-echelon simulation that propagates inventory dynamics across tiers would capture these cascading effects and identify system-level optima that may differ from the material-level recommendations produced by the current approach.

REFERENCES

Bendavid, I., Herer, Y. T., & Yücesan, E. (2017). Inventory management under working capital constraints. *Journal of the Operational Research Society*, *68*(10), 1145–1156. https://doi.org/10.1057/s41274-016-0131-4

Chen, L., Dong, T., Peng, J., & Ralescu, D. (2023). Uncertainty Analysis and Optimization Modeling with Application to Supply Chain Management: A Systematic Review. *Mathematics*, *11*(11), 2530. https://doi.org/10.3390/math11112530

Chopra, S., & Meindl, P. (2013). *Supply Chain Management: Strategy, Planning, and Operation* (5th ed.). Pearson.

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. J. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, *6*(1), 3–33.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, *74*(366), 427–431.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. https://otexts.com/fpp3/

Künsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics*, *17*(3), 1217–1261. https://doi.org/10.1214/aos/1176347265

Law, A. M., & Kelton, W. D. (2007). *Simulation modeling and analysis* (Vol. 3). New York: Mcgraw-hill.

Ljung, G. M., & Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika*, *65*(2), 297–303.

Muller, M. (2011). *Essentials of Inventory Management* (2nd ed). AMACOM.

Newell Brands. (2026). *2025 Annual Report*. https://ir.newellbrands.com/static-files/158c1f12-4590-49d1-a7c7-8e1be14343ae

Taskin, A. (2009). A Multi-Echelon Inventory Management Framework for Stochastic and Fuzzy Supply Chains. Expert Systems with Applications. https://doi.org/10.1016/J.ESWA.2008.06.082