Proceedings of the 11th North American International Conference on Industrial Engineering and Operations Management, Milwaukee, Wisconsin, USA, June 9-11, 2026

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/NA11.20260242
Published: June 9, 2026

Stress-Conditioned Monte Carlo Modeling of Stockout Risk on the Savannah-Atlanta Lane

Neil Sharma and Aditya Manchanda
Independent Researchers
Alpharetta, GA, USA
neilsharma.research1@gmail.com

Abstract

Congestion at major container ports can cause delivery delays, increasing the risk of grocery stockouts further downstream. This paper examines how congestion at the Port of Savannah affects the reliability of shipments on the Savannah–Atlanta replenishment corridor. To study this, a stress-conditioned Monte Carlo inventory model that links a public container throughput proxy (TEU volume) to discrete congestion regimes $R \in \{-1,0,1,2\}$, representing low through extreme stress was built. Each stress level alters the distribution of shipment delays, including the probability of rare but severe disruptions. The model simulates how stochastic lead times translate into inventory depletion and estimates the stockout probability $P(\text{stockout}|R,B,s)$ and the expected shortage across inventory buffer levels $B \in \{2,3,4,5,6\}$ days of coverage and policy scenarios $s \in$ {Baseline, Safety Stock (+1 day), Tail Mitigation}. Results show that congestion shocks significantly increase the risk of stockouts when buffers are small. Under Extreme stress ($R=2$) with a low buffer ($B=2$), baseline stockout risk reaches 12.51%, compared with 1.83% under a one-day safety stock increase and 9.31% under tail-focused mitigation. Shortage severity follows the same pattern: expected shortage falls from 6.370 units in the baseline scenario to 0.849 units with safety stock. When buffers increase to $B=3$, baseline stockout risk drops to 1.76%, and further to 0.24% with a safety stock. Overall, the results suggest that increasing buffer coverage is the strongest lever for reducing both the frequency and magnitude of stockouts, while mitigation strategies that reduce extreme delays become more effective once buffer levels are no longer critically low. The framework shows how upstream port congestion indicators can be translated into downstream inventory risk estimates to support practical supply chain planning decisions.

1. Introduction and Project Framing

1.1 Context

This report examines grocery replenishment reliability on the Savannah, Georgia-to-Atlanta, Georgia lane. When the system is under stress, deliveries can arrive later than planned. In inventory planning, the most damaging problems usually come from larger, less frequent delays rather than small day-to-day variation. A few long delays can push lead time beyond the inventory buffer and cause stockouts, even if the average delay appears acceptable. Because shipment-level delay data for private supply chains are typically not publicly available, we use a public port activity series as a consistent indicator of system stress and classify time into four stress regimes. We then simulate stress-conditioned delays and track how often inventory runs out and how large the shortages become. The key idea in this paper is to allow delay risk, especially rare severe delays, to change with system stress rather than assuming a fixed delay distribution. This creates a stress-conditioned inventory risk model that links upstream congestion to downstream stockout risk. The results are then translated into simple decision rules for buffer sizing under uncertainty using only public data.

1.2 Decision Question

Decision: choose an inventory buffer level B and a mitigation strategy for the Savannah to Atlanta replenishment lane (Baseline, Safety stock, or Tail mitigation).
Objective: keep P(stockout) at or below 1 percent as the primary target and at or below 2 percent as a secondary target under each stress regime R.

Control variables and notation used in this report:
R: stress regime index derived from the stress proxy.
B: days of cover inventory buffer.
s: scenario or policy option (Baseline, Safety stock, Tail mitigation).
P(stockout | R, B, s): stockout probability under the specified conditions.
E(shortage | R, B, s): expected shortage severity under the specified conditions.

Executive Summary
This report evaluates how stress-driven delays increase the risk of grocery stockouts on the Savannah, Georgia, to Atlanta, Georgia, replenishment lane. It quantifies how inventory buffers and mitigation strategies reduce that risk. We use a stress-conditioned Monte Carlo inventory model to estimate stockout probability and shortage severity across stress regimes, days-of-cover, and policy scenarios.

2.1 Problem Statement

Congestion-driven delivery delays increase the risk of grocery stockouts on the Savannah, Georgia-to-Atlanta, Georgia, replenishment lane, especially under High and Extreme stress. This affects planners, stores, and customers through service failures and higher operating costs.

2.2 Model Summary (what we did)

A public TEU-based stress signal is mapped to stress regimes (**R = −1, 0, 1, 2**). Stress increases the likelihood of severe delays, delays increase lead time, and longer lead time increases the chance that inventory is depleted before replenishment arrives. We evaluate outcomes using **P(stockout | R, B, s)** and expected shortage for buffer levels **B = 2–6 days** and scenarios **s ∈ {Baseline, Safety stock (+1 day), Tail mitigation}**.

2.3 Key Quantified Findings

1) Under Extreme stress (R=2) and low buffer (B=2 days), baseline stockout risk is high: P(stockout)=12.51%. Safety stock reduces this to 1.83%, while Tail mitigation reduces it to 9.31%.
   Implication: when the buffer is very low under Extreme stress, adding buffer is the fastest and most effective lever.
2) Under Extreme stress (R=2) at B=3 days, baseline risk is 1.76%, Safety stock is 0.24%, and Tail mitigation is 1.22%.
   Implication: one additional day of buffer via Safety stock can move operations below the 1% service target at B=3 under Extreme stress.
3) Severity follows the same pattern. Under Extreme stress (R=2) at B=2 days, the expected shortage is 6.370 units at baseline, 0.848945 units with Safety stock, and 4.656 units with Tail mitigation.
   Implication: buffer-based actions reduce not only the frequency of failures but also the severity of shortages when they occur.
4) Decision rules for the 1% service target (from Table 1). Under High stress (R=1), baseline requires B≥4, Safety stock requires B≥2, and Tail mitigation requires B≥3. Under Extreme stress (R=2), baseline requires B≥4, Safety stock requires B≥3, and Tail mitigation requires B≥4.

2.4 Recommendation

For Extreme-stress resilience using P(stockout) ≤ 1.0%, operate at baseline B=4 days, or operate at B=3 days when Safety stock (+1 day) is available; Tail mitigation improves outcomes but does not replace a sufficient buffer at low B.

1) Background and Stakeholders

3.1 Context and Why Lead Time Reliability Matters

This project focuses on grocery replenishment on the Savannah-to-Atlanta lane. In this setting, the key operational issue is lead time reliability, not just average lead time. When deliveries arrive late, inventory must cover demand for longer than planned. The most damaging outcomes are usually driven by less frequent but severe delays because those events can push lead time beyond the inventory buffer and cause stockouts. For this reason, resilience planning must account for how system stress changes the likelihood of severe delays.
Implication: planning must protect against severe delays, not just average lead time.

3.2 Decision Maker Framing and Control levers

The primary stakeholder is a distribution or inventory planner who sets replenishment buffers and service targets for this lane. The planner's objective is to keep stockout risk below a target threshold under different stress conditions. The main control levers are the inventory buffer B and the mitigation scenario, S. Performance is evaluated using P(stockout | R, B, s) as the main frequency metric and E(shortage | R, B, s) as the main severity metric. The primary service target is P(stockout | R, B, s) ≤ 1%, with ≤ 2% as a secondary threshold.
Implication: the planner's decision is to choose B and s so the service target is met for each stress regime R.

4. Data Sources and Processing

4.1 Public Data Source and What it Represents

We use a public monthly time series of Port of Savannah throughput reported in TEUs (twenty-foot equivalent units). The TEU series is used only as a **stress proxy** for system load that can plausibly increase delay risk in downstream freight movement. We do not treat TEU throughput as a direct measure of truck delays on the Savannah-to-Atlanta replenishment lane. Instead, TEU is used to classify time periods into stress regimes that condition the delay model.
Source: GPA Market Intelligence, "Port of Savannah – TEU Total for All Terminals," monthly TEU totals through December 2025.

4.2 Standardization: TEU to a Stress Index

To compare months on a common scale, we standardize TEU throughput to a z-score. Let **μ_TEU** and **σ_TEU** be the mean and standard deviation of **TEU_t** computed over the full monthly window **Jan 2003–Dec 2025**. We compute:

$$z_t = \frac{TEU_t - \mu_{TEU}}{\sigma_{TEU}}$$

This produces a dimensionless stress index, with positive values indicating above-average throughput and negative values indicating below-average throughput.
Implication: the standardized index supports a consistent, auditable rule for stress classification across the full historical window.

4.3 Mapping the Stress Index to Stress Regime R

We map the TEU z score **z_t** into a discrete stress regime **R ∈ {−1, 0, 1, 2}**, interpreted as Low, Normal, High, and Extreme stress, using fixed thresholds:

- R = −1 (Low) if $z_t \le a$
- R = 0 (Normal) if $a < z_t \le b$
- R = 1 (High) if $b < z_t \le c$
- R = 2 (Extreme) if $z_t > c$

with cutoffs (rounded to 4 decimals):

- $a = -0.5016$
- $b = 0.4981$
- $c = 1.4989$

We use the stress regime R as the conditioning variable for the delay model: higher R increases the probability of a severe delay, which in turn increases lead-time exposure and therefore increases **P(stockout | R, B, s)** and **E(shortage | R, B, s)**.

4.4 Data Handling and Reproducibility Checks

The TEU series is used as a complete monthly sequence from **Jan 2003–December 2025**, with no smoothing or seasonal adjustment applied. The stress index and regime assignment are produced programmatically. Monte Carlo results in later sections use **N = 50,000** simulations per **(R, B,s**) and a fixed seed of **42**, with **B ∈ {2,3,4,5,6}** and **R ∈ {−1,0,1,2}**.

4.5 Data Summary and Actuarial Data Categories

This report uses one public data source: monthly Port of Savannah TEU throughput from Jan 2003 through Dec 2025 (276 monthly observations). The series is complete at a monthly frequency with no missing months, and we apply no smoothing or seasonal adjustment. All processing steps (z-score and regime mapping) are reproducible from the stated rules.

In the Actuarial Process Guide framing, our inputs and outputs align with three important data areas. TEU and its z-score define operating conditions by assigning Low, Normal, High, and Extreme stress regimes. The delay mixture model defines the range and severity of disruption outcomes (including tail behavior). The simulation outputs quantify likelihood and loss metrics by reporting **P(stockout | R, B, s)** (frequency) and expected shortage (severity) across stress regimes and scenarios.

5. Model Overview

5.1 System Flow

Figure 1 shows the model in one view. We start with the public monthly TEU series and turn it into a stress score. That score is grouped into a stress regime: Low, Normal, High, or Extreme. The stress regime controls how likely a severe delay is. We then draw a delay type (short, moderate, or severe) and a delay amount in hours. The delay is added to the base lead time to get the total lead time. During that lead time, demand occurs, and starting inventory is set based on the days-of-cover buffer **B**. If demand during the lead time exceeds available inventory, a stockout occurs, and a shortage is recorded. Repeating this simulation many times produces the main decision metrics, including **P(stockout | R, B, s)** and expected shortage, for each stress regime, buffer level, and scenario.

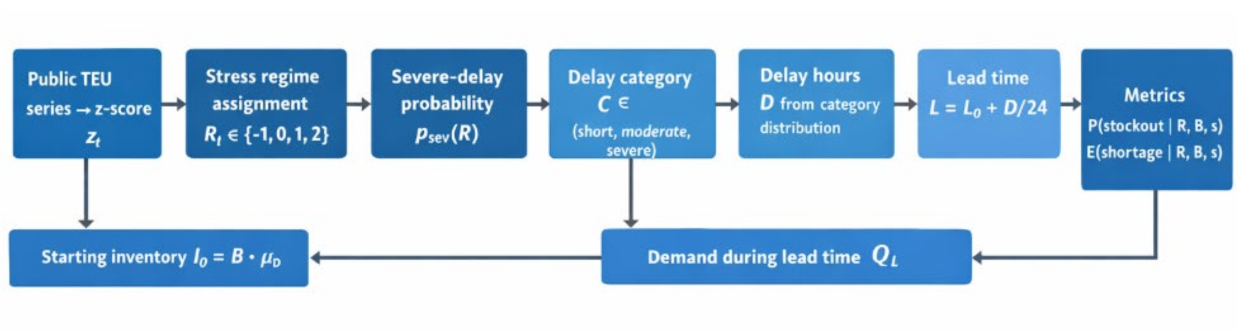

Figure 1. Model flow diagram

5.2 Variables and Notation (quick reference)

- **Stress regime R:** discrete stress level derived from the TEU z-score, with Low (R = -1), Normal (R = 0), High (R = 1), and Extreme (R = 2).
- **Days-of-cover B:** planned inventory buffer measured in days of average demand.
- **Scenario s:** policy options, s ∈ {Baseline, Safety stock, Tail mitigation}.
- **Base lead time L0:** nominal lead time in days before adding delay; L0 = 1.0.
- **Delay D:** additional delay in hours generated by the delay mixture model.
- **Lead time L:** total lead time in days, L = L0 + D/24.
- **Mean daily demand μD:** average demand per day; μD = 100.
- **Demand variability σD:** demand standard deviation per day; σD = 25.
- **Stockout:** the event that demand during lead time exceeds the starting inventory
- **Shortage:** unmet demand during the lead time when inventory is insufficient.
- **Primary metrics: P(stockout | R, B, s)** and expected shortage.

5.3 Scenarios (policy levers)

This report evaluates three scenarios **s**, which correspond to practical policy choices.
Baseline: uses the stress-conditioned delay and demand models as specified.
Safety stock: increases the buffer by one day, implemented as $B' = B + k$ with $k = 1$. This changes the starting inventory while leaving the delay process unchanged.
Tail mitigation: reduces the risk of severe delay by reducing stress sensitivity in the severe-delay probability mapping. In the implemented configuration, this is $\beta \leftarrow 0.8\beta$. This changes the delay process while leaving the buffer definition unchanged.
One important distinction: Safety stock changes the amount of inventory available, while Tail mitigation changes the delay process that consumes that inventory.

5.4 What the Model Produces

For each stress regime **R**, buffer level **B**, and scenario **s**, the simulation estimates the frequency and severity of service failure. Frequency is summarized by **P(stockout | R, B, s)**. Severity is summarized by expected shortage. Because severe delays drive the largest failures, we also report tail metrics for delays and lead time (for example, P95 and exceedance probabilities such as P(delay > 12 hours) and P(delay > 24 hours)).

6. Mathematical Methodology

6.1 Stress to Severe-Delay Probability Mapping

We model stress as a driver of severe delays rather than assuming a fixed delay profile in every month. For a given stress regime **R**, we define the probability that a delay is severe using a logistic mapping. We use a logistic mapping, which is a standard probability-modeling approach, to relate stress regime to severe-delay probability:

$$p_{sev}(R) = \frac{1}{1+\exp\left(-(\alpha+\beta R)\right)}$$

where **α=−2.197** and **β=0.811.** Here, **R** is the stress regime, **α** is the intercept, and **β** controls how strongly severe-delay probability rises with stress. The logistic component increases with **R**, so higher stress regimes yield a higher probability of severe delay. We chose **β=0.811** so that severe-delay probability approximately doubles between Normal (**R = 0**) and High stress (**R = 1**). This calibration is directionally consistent with freight reliability measures based on **95th-percentile travel time**, such as the **Planning Time Index (PTI)** used by FHWA, and with reported PTI values above **1.4** on major freight bottlenecks. In this model, **γ** is introduced in Section 6.2 as a fixed mixing weight that allocates the remaining non-severe probability mass between short and moderate delays.

6.2 Allocation of Non-Severe Probability: Short vs Moderate

Conditional on a delay being non-severe, we split probability mass between short and moderate outcomes using the fixed mixing parameter γ:

$$p_{\text{mod}}(\mathbf{R}) = \boldsymbol{\gamma}\,(1-p_{\text{sev}}(\mathbf{R})\,)$$

$$p_{\text{short}}(\mathbf{R}) = (1-\boldsymbol{\gamma})\,(1-p_{\text{sev}}(\mathbf{R})\,)$$

With $\gamma = 0.35$, this means 35% of the non-severe probability is assigned to the moderate delay component and 65% to the short delay component.
This allocation uses a standard mixture-model formulation, where a fixed weight divides probability mass across components

6.3 Delay Severity Components and Distributions

Given regime R, each simulation samples one delay realization **D** (in hours) from a three-component mixture:

Short delay: D_short ~ Triangular(**a_s, m_s, b_s)**
Moderate delay: D_mod ~ Triangular(**a_m, m_m, b_m**)
Severe delay: D_sev ~ LogNormal(median = **d̃**, sigma = **σ_sev**), optionally capped at **cap_hours (72 hours)**
Short and moderate delays represent routine operational variation, such as minor scheduling changes, small congestion effects, or loading delays. Severe delays represent less frequent but operationally significant disruptions, such as port congestion, equipment shortages, or cascading logistics delays.

The delay components use standard distributional choices: triangular distributions for routine variation and a lognormal distribution for rare severe delays, which is appropriate for right-skewed delay behavior. The severe-delay distribution uses **σsev=1.2\sigma_{sev} = 1.2σsev=1.2**, which yields upper-tail delays of about **60 hours** at the 90th percentile. This value is used as an assumed heavy-tail calibration intended to represent rare but operationally significant disruptions rather than a direct estimate from lane-level shipment data.

To prevent unrealistically long simulated outcomes, severe delays are capped at **72 hours**. This reflects a practical operational ceiling consistent with trucking constraints and typical intervention points in freight operations. This three-category structure allows the simulation to capture both routine variation and rare, extreme disruptions, which is important because stockout risk is driven mainly by these longer-tail events

6.4 Mixture Sampling Procedure (One Trial)

For each trial:

1. Compute **p_sev(R)** from Section 6.1

2. Compute **p_mod(R)** and **p_short(R)** from Section 6.2

3. Draw **u ~** Uniform(0, 1) and select the component
   If **u ≤ p_sev(R)**: sample **D = D_sev**
   Else if **u ≤ p_sev(R) + p_mod(R):** sample **D = D_mod**
   Else: sample **D = D_short**

So R controls how often the system enters the severe tail, and γ controls how the remaining "normal operations" probability is split between short and moderate outcomes.

6.5 Lead Time Construction and Fractional Day Handling

We represent lead time in days as a base lead time plus a stochastic delay:

$$\mathbf{L} = \mathbf{L_0} + \frac{\mathbf{\Delta}}{24}$$

where **L0** is baseline lead time (days), and **D** is the sampled delay (hours). Because **D/24** can be fractional, **L** is treated as continuous rather than rounded to whole days.

6.6 Demand During Lead Time (Aggregate Over a Continuous Window)

After sampling **L**, we simulate total demand accrued over that lead time window as a single aggregate draw:

$$D_L \sim \mathcal{N}(\mu_D L,\ \sigma_D^2 L)$$

where **D_L** is total demand during lead time, **μ_D** is mean daily demand, **σ_D** is daily demand standard deviation, and **L** is realized lead time in days. This formulation follows the standard aggregation properties of stochastic demand over a continuous-time window: the mean scales with **L,** the variance scales with **L**, so the standard deviation scales with √**L**. It preserves fractional-day effects without discretizing demand into **ceil(L)** daily steps. This follows the standard aggregation properties of stochastic demand over a continuous-time window.

6.7 Interpretation and Relevance to the Risk Metrics

This delay model is the mechanism that links upstream port stress to downstream service risk. In each trial, the stress regime **R** controls the mixture weights through **p_sev(R),** which governs how frequently the process enters the severe-delay tail. When a severe component is selected, sampled delay **D** increases the realized lead time **L = L0 + D/24**, which expands the demand exposure window and raises aggregate lead-time demand **D_L**. Because stockout and shortage are evaluated by comparing **D_L** against fixed on-hand inventory **I** (set by **B** days of cover), the primary driver of adverse outcomes is tail behavior: higher **R** increases the probability of severe delays, which increases the upper tail of **L** and therefore increases **P(stockout)** and **E(shortage)**, especially at low **B.** Practically, this section justifies why "tail-focused" mitigation (reducing tail probability or tail magnitude) can materially improve performance under high and extreme regimes, while additional buffer (safety stock) is most effective at low days-of-cover because it directly increases **I** relative to uncertain **D_L.**

7. Inventory Depletion Risk Model (Lead Time → Stockout Risk)

7.1 Purpose

This section converts stochastic lead time into decision-relevant inventory risk. Given a realized replenishment lead time **L**, the system faces uncertain demand over that lead-time window. A stockout occurs when cumulative demand during **L** exceeds on-hand inventory I. We report both the frequency of stockouts and the severity of shortages to align with actuarial frequency–severity framing.

7.2 Inventory Policy Parameterization (days-of-cover)

We represent the inventory buffer using **the days-of-cover metric B**, which is intuitive to nontechnical decision-makers. Let **μ_D** denote the mean daily demand (units/day). We define starting inventory on hand, **I** as:

$$\mathbf{I = \mu D \cdot B.}$$

Where **I** is the starting inventory on hand, **μD** is the mean daily demand, and **B** is the days of cover
In mitigation scenarios, the effective days of cover become:

- Baseline: **B_eff = B**
- Safety stock: **B_eff = B + k (** where **k = 1** day in our configuration)
- Tail mitigation: **B_eff = B** (inventory unchanged; delay tail modified upstream)

7.3 Realized lead time

Each Monte Carlo trial produces a delay draw **D** (hours) from the stress-conditioned delay model (Section 6). We convert that delay into realized lead time in days:

$$\mathbf{L = L0 + D/24,}$$

where **L0** is baseline lead time (days). **L** is treated as continuous (fractional days allowed)

7.4 Aggregate Demand During Lead Time

Let **D_L** denote total demand during lead time **L**. We model demand during the lead-time window as an aggregate draw whose mean and variability scale with the window length:

$$D_L \sim \mathcal{N}(\mu_D L,\ \sigma_D^2 L)$$

Where **DL** is the total demand during the lead time, **μD** is the mean daily demand,**σD** is the daily demand standard deviation, and **L** is the lead time in days. Equivivalently,

$$E[D_L] = \mu_D L, \qquad SD(D_L) = \sigma_D \sqrt{L}.$$

This approach preserves fractional-day effects without rounding **L** up to whole days, reflecting the operational reality that longer replenishment windows create more opportunity for demand to exceed expectations.

7.5 Stockout Event Definition and Shortage Magnitude

Given the on-hand inventory **I** and realized lead-time demand **D_L**, we compute:

Stockout indicator

$$\mathbf{S} = \mathbf{1}[\mathbf{D_L} > \mathbf{I}].$$

Shortage magnitude

$$\mathbf{Sh} = (\mathbf{D_L} - \mathbf{I})^+ = \max(\mathbf{D_L} - \mathbf{I}, 0).$$

This yields both:

- Frequency (whether a stockout occurred)
- Severity (how large the unmet demand is when it occurs).

This follows a standard inventory risk formulation in which a stockout occurs when lead-time demand exceeds available inventory.

7.6 Reported Inventory Risk Metrics

Across **N** Monte Carlo trials per configuration (**R**, **B**, scenario), we report:

- **Stockout probability**

$$\mathbf{P_{stockout}} = \mathbb{E}[\mathbf{S}]$$

(empirically: the sample mean of **S** across trials)

- **Expected shortage**

$$\mathbf{E_{shortage}} = \mathbb{E}[\mathbf{Sh}]$$

We also retain tail diagnostics from the lead-time layer (Section 6) because they explain the mechanism behind elevated stockout risk:

- **Delay_P95**, **LeadTime_P95**
- **P(D > 12)**, **P(D > 24)**

Uncertainty is quantified using confidence intervals:

- **P_stockout** uses a binomial confidence interval method (e.g., the Wilson method).
- **E_shortage** uses a bootstrap confidence interval for the mean (when enabled).

7.7 Interpretation

Stockouts occur when lead-time demand exceeds the inventory buffer. We set the buffer as **I = μ_D · B_eff** and compute realized lead time as **L = L0 + D/24**. Because **L** increases continuously with delay **D**, even sub-day disruptions expand the demand exposure window and increase aggregate lead-time demand **D_L**.

The results show a "knee" in **B** because **B** shifts the threshold **I**. At low **B**, **I** lies near the upper tail of **D_L**, so relatively small changes in lead time or demand variability cause frequent threshold crossings. As **B** increases, **I** moves right; once it clears most of the **D_L** tail, **P(stockout)** collapses quickly toward zero, and expected shortage becomes negligible.

Stress primarily increases risk through tail behavior. Higher stress **R** increases severe-delay probability **p_sev(R)**, increasing the frequency of large delays and shifting the right tail of **D** upward. This pushes the right tail of **L** upward, which in turn pushes the tail of **D_L** upward. Operationally, that increases **P(stockout)** (more crossings of **D_L > I**) and **E(shortage)** (bigger exceedances when crossings occur), especially in the low-to-moderate **B** region where the system is closest to the threshold.

The mitigation scenarios act through distinct mechanisms, which is why their effects differ across regimes. **Safety stock** increases **B_eff = B + k**, raising **I** for every trial, so it is most effective near the knee, where additional buffer changes whether the threshold is crossed. **Tail mitigation** leaves **I** unchanged but reduces extreme lead-time outcomes by shrinking the tail of **D/L**, making it most valuable under higher stress, where severe tail events dominate the risk. The decision rules in the Results section quantify these tradeoffs by identifying the smallest **B** that meets a service target (e.g., **P(stockout) ≤ 1%**) under each stress regime.

8. Results

8.1 Stress Increases the Chance of Severe Delays

Figure 2 shows the model's main mechanism: as the stress regime R increases, the probability of a severe delay rises sharply. Because inventory systems usually fail when rare long delays extend lead time beyond the buffer, this increase in the probability of severe delays directly increases stockout risk. In Figure 2, severe-delay probability rises from about 5% at R = -1 to 10% at R = 0, 20% at R = 1, and 36% at R = 2. This explains why a buffer that performs adequately under Normal conditions can fail under High or Extreme stress.

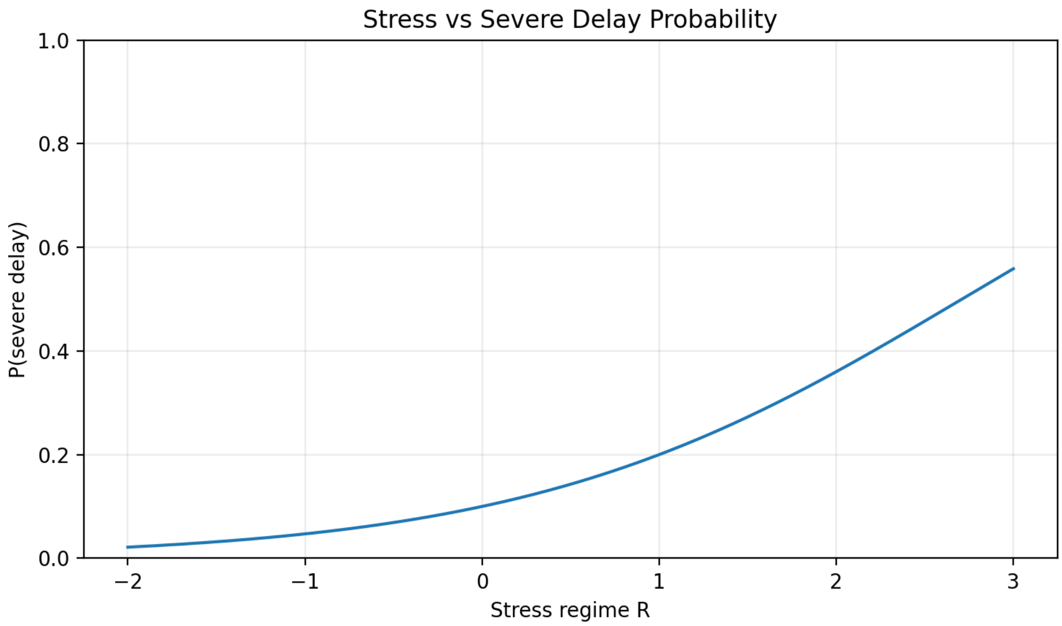

Figure 2. Stress vs Severe Delay Probability. The severe-delay probability increases with the stress regime R, providing the causal link between stress and tail risk.

8.2 Tail Mitigation Reduces Exposure to the Worst Delays

Figure 3 focuses on the upper tail of delays under High (**R = 1**) and Extreme (**R = 2**) stress. The key takeaway is simple: tail mitigation improves worst-case delay behavior compared to baseline in both regimes. That is exactly what tail mitigation is designed to do, and the plot shows it moving the delay tail in the desired direction.

The shift is not small. At **R = 1**, the P95 delay falls from roughly 26 hours in baseline to roughly **23 hours** under tail mitigation. At **R = 2**, the P95 delay falls from roughly **33 hours** to **about 29 hours**. Reducing the upper tail matters because stockouts are driven by these long-delay events, not by typical days.

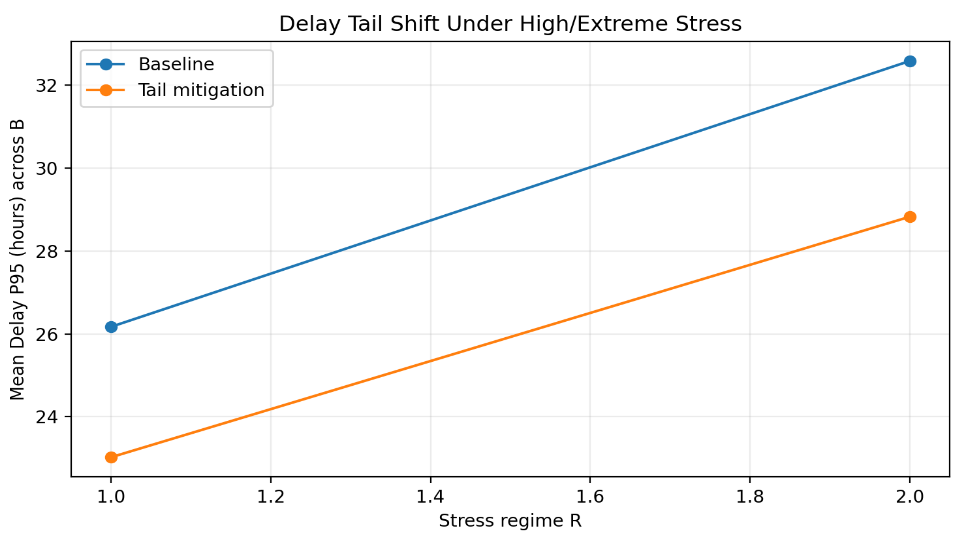

Figure 3. Delay Tail Shift Under High/Extreme Stress. Baseline vs tail mitigation for upper-tail delay behavior under R=1 and R=2

Impact: Tail mitigation reduces how often lead time becomes unusually long, which lowers stockout risk most when the buffer is moderate. At a very low buffer, adding inventory tends to dominate; at a very high buffer, most scenarios are already safe, and the marginal benefit shrinks.

8.3 Stockout Probability vs Days-of-Cover Team #23968 "Modeling Stockout Risk Under Stress."

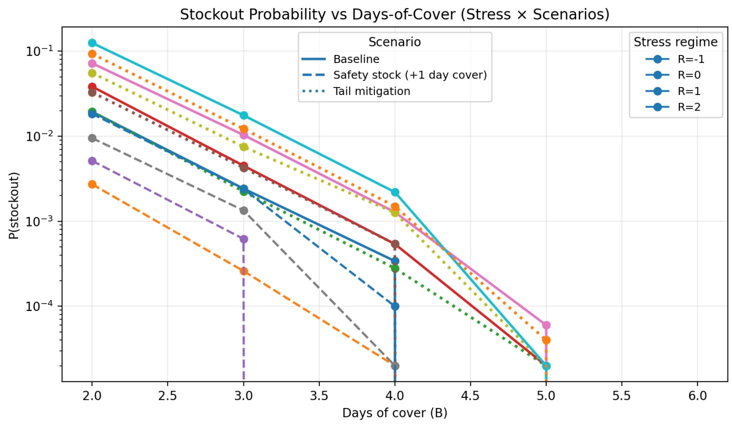

Figure 4. Stockout Probability vs Days-of-Cover (Stress x Scenarios)

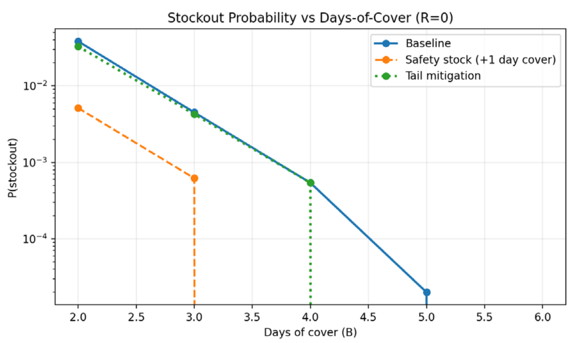

Figure 4a. Stockout vs Days-of-Cover (R=1)

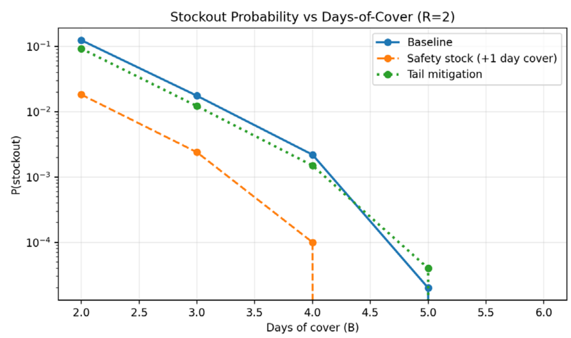

Figure 4b. Stockout vs Days-of-Cover (R=2)

Figure 4 is the main decision plot: as days-of-cover **B** increases, stockout probability drops sharply, and the largest improvements occur at low B. Safety stock has the strongest effect at low B because it directly increases the buffer, while tail mitigation reduces risk by reducing the frequency of severe delays.

Under Extreme stress (**R=2**) at **B=2**, baseline **P(stockout)=12.51%**; safety stock reduces it to **1.83%**, and tail mitigation reduces it to **9.31%**.
At **B=3**, the baseline is **1.76%**, safety stock is **0.24%**, and tail mitigation is **1.22%**.
This shows a practical cliff: moving from **B=2** to **B=3** is a major improvement, and safety stock can push risk below 1% at **B=3** under **R=2**.

Under High stress (**R=1**) at **B=2**, the baseline is **7.21%**; the safety stock is **0.95%**; the tail mitigation is **5.55%**.
The pattern is the same: a low buffer is fragile under stress, and adding a buffer is the fastest way to control risk.

8.4 Expected Shortage vs Days-of-Cover: How Severe Failures Are

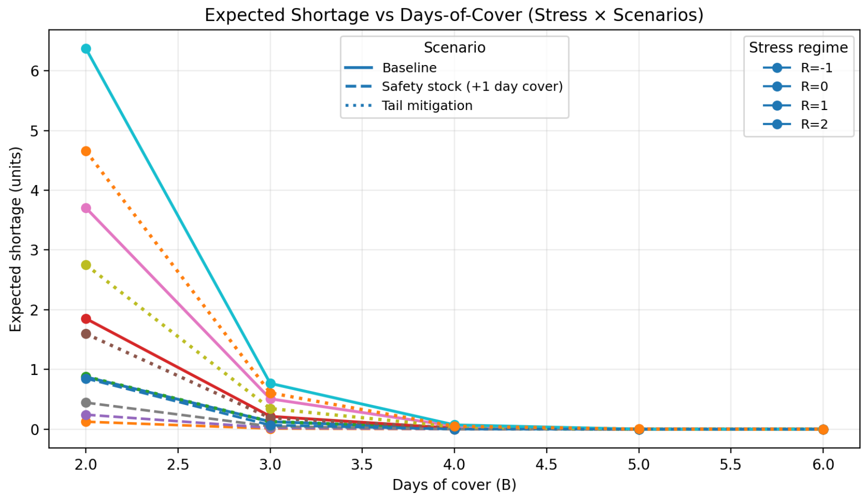

Figure 5. Expected Shortage vs Days-of -Cover

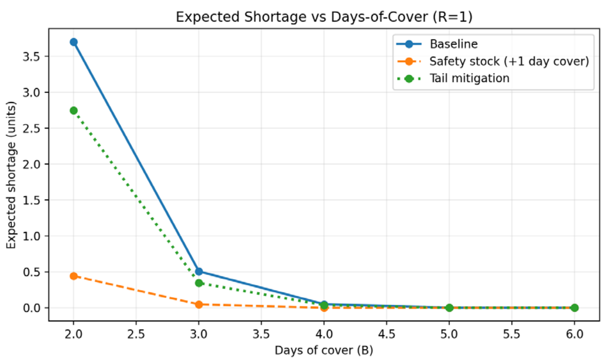

Figure 5a. Expected Shortage vs Days-of-Cover (R=1)

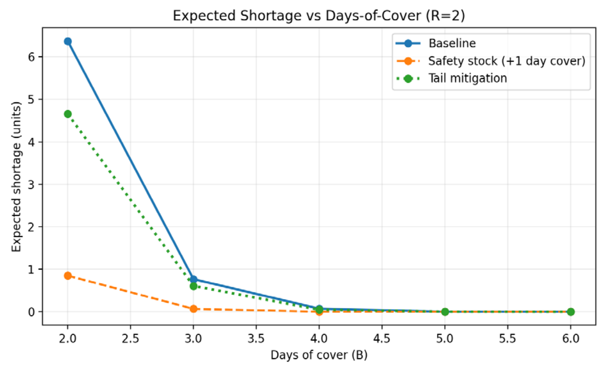

Figure 5b. Expected Shortage vs Days-of-Cover (R=2)

Expected shortage measures the average shortfall in **demand units** when inventory is not enough during the lead time. Figure 5 shows severity falls quickly as **B (days)** increases, and is highest under High and Extreme stress at a low buffer. Safety stock reduces severity the most at low **B** because it increases inventory directly. Because it averages the shortfall across trials, the expected shortage serves as an expected loss proxy in demand units, complementing **P(stockout | R, B, s)** as the frequency metric.

Under Extreme stress (**R=2**) at **B=2 days**, baseline expected shortage is **6.370 units**; safety stock is **0.848945 units**; tail mitigation is **4.656 units**.
Under High stress (**R=1**) at **B=2 days**, baseline is **3.703 units**; safety stock is **0.444163 units**; tail mitigation is **2.750 units**.
This supports the same takeaway as stockout probability: low-buffer operation under stress produces larger service failures, and buffer-based policies reduce both frequency and severity.

8.5 Decision Rules (Service Targets)

This section turns the plots into a simple rule a planner can use. Days-of-cover **B** is the number of days the system can continue selling at average demand if the next replenishment arrives late. A larger **B** gives more protection, but it also means holding more inventory. We evaluate each policy by whether it keeps the chance of a stockout below a target.

We use a primary service target of **P(stockout | R, B, s) ≤ 1.0%**, meaning stockouts occur in at most 1 out of 100 simulation trials for that stress level and policy. We also report a secondary target of **P(stockout | R, B, s) ≤ 2.0%**, which is less strict and can be useful when a planner is willing to accept slightly more risk. The decision rules are stated for High (**R=1**) and Extreme (**R=2**) stress because those are the conditions where buffer choice and mitigations matter most. Low- and Normal-stress results are still shown in the full curves and appendix tables as a baseline check.

Stress regime R	Scenario s	Min B (days) for P(stockout) ≤ 1.0%	Min B (days) for P(stockout) ≤ 2.0%
R=1 (High)	Baseline	**4**	**3**
R=1 (High)	Safety stock (+1 day)	**2**	**2**
R=1 (High)	Tail mitigation	**3**	**3**
R=2 (Extreme)	Baseline	**4**	**3**
R=2 (Extreme)	Safety stock (+1 day)	**3**	**2**
R=2 (Extreme)	Tail mitigation	**4**	**3**

Table 1. Minimum days-of-cover B (days) required to meet stockout-risk targets (by stress regime and scenario)

To read the table, choose the stress regime and scenario, then take the smallest **B** that meets the target. Under Extreme stress, baseline operation needs **B=4 days** to meet the 1% target, while safety stock allows the 1% target at **B=3 days**. Tail mitigation reduces risk, but in this grid, it does not reduce the minimum buffer requirement below **B=4 days** for the 1% target under Extreme stress. Under High stress, the same pattern holds, but the minimum buffers are slightly lower because the system is less stressed.

Impact: This table is the main deliverable because it converts all curves into a direct buffer rule keyed to stress regime and scenario.

8.6 Historical Credibility Anchor (Stress Aligns with Modeled Risk Over Time)

Figure 6 provides a credibility check for the stress proxy by comparing the **TEU z-score** with the modeled baseline stockout risk at a fixed buffer of **B = 3** days. Holding **B** fixed isolates the effect of stress over time. The two series move together: when the stress proxy rises, modeled stockout risk rises, and when stress falls, modeled risk falls. During the shaded congestion-era period, baseline risk at **B = 3** rises to about **1.9%** at peak, compared with about **0.5%–1.0%** in lower-stress periods outside that interval. This does not validate the model against observed stockout logs, but it shows that the stress proxy yields a stable, interpretable relationship with modeled risk over time.

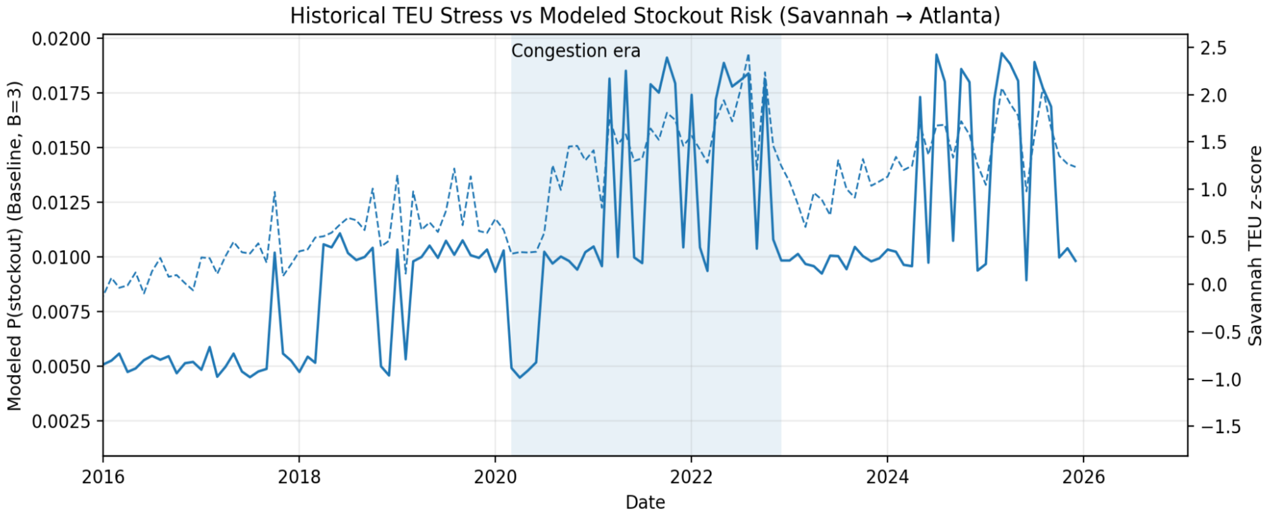

Figure 6. Historical TEU Stress vs Modeled Stockout Risk (Baseline, B=3 days). TEU z-score (unitless, dashed) and modeled baseline stockout probability (% of trials, solid) move together over time; shaded region highlights the congestion era.

8.7 Sensitivity Analysis on Key Parameters

β, σ_sev, the severe-delay cap, and **σ_D** are set by assumption rather than estimated from observed data. Because the buffer thresholds in Table 2 depend on these parameters, it is worth checking how much they change when the assumptions shift. Table 2 shows the minimum buffer required to meet each service target when each parameter is adjusted 20% up and 20% down, one at a time, holding everything else fixed.

The **R=1,** 2% target is the most stable and remains at **B=3** across all tested combinations. The other three columns are stable in most cases but shift by one day under certain tail assumptions. The **R=2**, 1% target drops from **B=4** to **B=3** when the severe-delay tail is made narrower (**σ_sev** × **0.8**), meaning the system needs less buffer if severe delays tend to be shorter than assumed. It does not rise above **B=4** for this target. The **R=1**, 1% and **R=2**, 2% targets each show a one-day increase under heavier-tailed or higher-stress-sensitivity settings (**β × 1.2** and **σ_sev × 1.2**).

The cap on severe delays and demand variability, **σ_D** have a smaller effect. Lowering the cap from 72 to 48 hours reduces the **R=1**, 1% requirement from **B=4** to **B=3**, which makes sense because a tighter cap limits how long the worst delays can be. Raising **σ_D** by 20% does not change any thresholds, indicating that the demand variability assumption has little influence on the buffer rules within the ranges tested.

Table 2. Minimum days-of-cover B (baseline scenarios)

Parameter setting	R=1 (High) Min B ≤ 1.0%	R=1 (High) Min B ≤ 2.0%	R=2 (Extreme) Min B ≤ 1.0%	R=2 (Extreme) Min B ≤ 2.0%
β × 0.8	3	3	4	3
β × 1.0 (base)	4	3	4	3
β × 1.2	4	3	4	4
σ_sev × 0.8	3	3	3	3
σ_sev × 1.0 (base)	4	3	4	3
σ_sev × 1.2	4	3	4	4
cap 48h	3	3	4	3
cap 72h (base)	4	3	4	3
cap 96h	4	3	4	3
σ_D × 0.8	3	3	4	3
σ_D × 1.0 (base)	4	3	4	3
σ_D × 1.2	4	3	4	3

Overall, the Table 2 thresholds are stable within one day across all perturbations. **σ_sev** is the most important parameter. When the severe-delay tail is narrower or wider than assumed, the minimum buffer shifts accordingly. This means the decision rules in Table 2 are reasonable under a range of assumptions, but they are not guaranteed to hold if the true tail behavior differs materially from the assumed value. Calibrating **σ_sev** to observed lane-level shipment data is therefore the highest-priority model improvement identified in Section 11 (Future Work).

8.8 Validation and Robustness Checks

Validation checks were applied across all combinations of stress regime, buffer level, and scenario to confirm that the model preserves the expected directional relationships: higher stress should increase risk, larger buffers should reduce stockout probability and expected shortage, Safety stock should not worsen outcomes relative to baseline, and Tail mitigation should improve or at least not materially worsen delay-related metrics. The model achieves a **94.7%** pass rate across these checks. The remaining failures occur only in near-zero-risk regions, where both **P(stockout)** and **E(shortage)** are already extremely small, so the reversals are attributable to Monte Carlo noise rather than a structural problem.

The threshold results are also stable and operationally interpretable. Under baseline, maintaining **P(stockout | R, B, s) ≤ 1%** requires **3 days** of cover in Low and Normal stress and **4 days** in High and Extreme stress. Safety stock achieves the same target with **2 days** of cover through **R = 1** and **3 days** at **R = 2**. Tail mitigation improves outcomes under higher stress by reducing exposure to extreme delays, but it is generally weaker than Safety stock as an independent lever. Overall, the model behaves consistently in the configurations that matter most for planning.

9. Recommendations

This section translates Section 8 into an operating policy a planner can apply. Recommendations use the primary service target **P(stockout | R, B, s) ≤ 1.0%** and the minimum-buffer thresholds in **Table 1 (Page 18)**.

9.1 Mitigation Categories Considered

We considered three mitigation categories from the Actuarial Process Guide. Modifying outcomes through inventory buffering (days-of-cover and Safety stock) and tail reduction (Tail mitigation) is directly modeled and quantified in this report. Behavior changes are also considered, including routing flexibility, tighter appointment discipline, priority handling during High and Extreme stress, and escalation triggers when stress shifts upward; these actions reduce the chance that disruptions compound into extreme delays. Insurance or contracting approaches such as expedited-capacity agreements, service-level guarantees, or financial coverage for disruption costs can transfer or reduce financial impact, but are not explicitly modeled here because contract terms and cost data are not available from public sources.

9.2 Operating Policy

Use **Table 1 (Page 18)** as the rulebook for High **(R=1)** and Extreme **(R=2)** stress. The simplest robust baseline policy is **B = 4** days, which meets the 1% target under Extreme stress in the baseline scenario. If Safety stock can be deployed, the system can operate at **B = 3** days under Extreme stress while still meeting the 1% target. Tail mitigation improves outcomes but does not reduce the minimum buffer requirement below the baseline Extreme-stress threshold at the 1% target.

9.3 How to Use the Mitigations

Safety stock is the best short-run lever when stress rises, and **B** is below the Table 1 threshold because it increases inventory immediately. Tail mitigation reduces exposure to the longest delays under High and Extreme stress, lowering risk without increasing inventory, but it cannot fully protect very small buffers because long lead times can still occur. In practice, Safety stock is the immediate-response lever, while Tail mitigation is a longer-horizon investment that reduces the frequency of extreme-delay events.

9.4 Tradeoffs and Implementation

In routine planning, set **B** using Table 1 for the expected stress regime and scenario. When stress shifts into High or Extreme regimes, compare the current **B** to the Table 1 minimum and deploy Safety stock or increase **B** if the target is not met. The main trade-off is that higher B increases inventory carrying and spoilage exposure, while tail mitigation requires operational investment to reduce the frequency of extreme delays.

10. Limitations

This project is designed to produce decision-useful buffer and mitigation rules under stress using public data and a transparent simulation model. The results should be interpreted as planning guidance under uncertainty rather than precise forecasts for any specific facility, retailer, or product category. First, TEU throughput is a stress proxy, not a direct measure of truck delays on the Savannah-to-Atlanta lane. It supports reproducibility, but it does not capture all drivers of delay or establish a direct causal link to lane-level travel time performance. Second, demand is stylized and not SKU-, store-, or category-specific, so expected shortage should be interpreted as a comparative severity metric across scenarios rather than a direct estimate of units lost for a specific operation. Third, the model is not validated against lane-level shipment tracking data or proprietary stockout logs; the historical plot is intended only as a credibility anchor showing that the stress proxy and modeled risk move together over time. Fourth, the lane representation is simplified and does not explicitly model factors such as routing changes, labor disruptions, weather, equipment constraints, or adaptive responses like expedited shipments and warehouse rebalancing. Because TEU is a proxy and demand is stylized, the minimum-buffer thresholds in Table 1 should be interpreted as conservative planning guidance rather than exact guarantees. Overall, the model is best used for stress planning and policy direction, with the understanding that its decision rules can be improved as additional data becomes available

11. Conclusion

This report analyzed grocery replenishment reliability on the Savannah, Georgia, to Atlanta, Georgia replenishment lane under varying stress conditions. The central finding is that stress increases the chance of severe delays, and those rare long delays are what most often push lead time beyond the inventory buffer. As a result, stockout risk is highly nonlinear in days-of-cover: small increases in B at low buffer levels can reduce P(stockout | R, B, s) by large amounts under High and Extreme stress.

The main decision output is the minimum-buffer rule set in Table 1 (Page 18). Using the primary target P(stockout | R, B, s) ≤ 1.0%, baseline operation under Extreme stress (R=2) requires B = 4 days, while B = 3 days can meet the target when Safety stock is available. Tail mitigation improves outcomes by reducing exposure to the longest delays, but it does not replace the need for sufficient days-of-cover at a low buffer.

Overall, the project provides a reproducible, public-data-based framework that converts a general resilience question into clear buffer and mitigation rules. The results are most useful for stress planning and policy direction: they show how much buffer is needed to maintain service under High and Extreme conditions and how different mitigation choices change both the likelihood and severity of service failures.

References

GPA Market Intelligence, *Monthly TEU Throughput (Port of Savannah): TEU Total for All Terminals*, Dec. 2025. Accessed Mar. 1, 2026.

Modeling the Future Challenge (MTFC), *2025–26 Project Phase Guide* (Actuarial Process Guide sections referenced for required report components and the "three important data areas").

E. B. Wilson, "Probable Inference, the Law of Succession, and Statistical Inference," *Journal of the American Statistical Association*, 1927.

B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, 1993.

U.S. Department of Transportation, Federal Highway Administration, "Freight Mobility Trends Instructions," FHWA Freight Management and Operations.

U.S. Department of Transportation, Federal Highway Administration, "2020 National List of Major Freight Highway Bottlenecks and Congested Corridors," FHWA Freight Management and Operations

U.S. Department of Transportation, Federal Highway Administration, *Travel Time Reliability: Making It There On Time, All The Time, 2026*.

Appendix: Verification Tables (High and Extreme Stress)

This page provides quick verification of the main Risk Analysis outputs used in Section 8 and the decision rules in Table 1. Frequency is reported as **P(stockout | R, B, s)** (percent of simulation trials) and severity as expected shortage (demand units). Delay tail behavior is summarized by delay P95 (hours) and P(delay > 24h) (percent of trials).

Table A1. Stockout frequency and shortage severity (R = 1 and 2; B = 2, 3, 4)

R	B (days)	P(stockout) baseline	P(stockout) safety	P(stockout) tail	E(shortage) baseline (units)	E(shortage) safety (units)	E(shortage) tail (units)
1	2	7.21% [6.99, 7.44]	0.95% [0.87, 1.04]	5.55% [5.35, 5.75]	3.703 [3.533, 3.874]	0.444 [0.390, 0.496]	2.750 [2.612, 2.896]
1	3	0.83% [0.75, 0.91]	0.08% [0.05, 0.11]	0.52% [0.46, 0.58]	0.304 [0.257, 0.350]	0.022 [0.012, 0.032]	0.180 [0.149, 0.211]
1	4	0.08% [0.05, 0.11]	0.00% [0.00, 0.01]	0.04% [0.02, 0.06]	0.019 [0.011, 0.027]	0.000 [0.000, 0.001]	0.010 [0.005, 0.015]
2	2	12.51% [12.22, 12.80]	1.83% [1.71, 1.95]	9.31% [9.06, 9.56]	6.370 [6.123, 6.617]	0.849 [0.785, 0.914]	4.656 [4.457, 4.855]
2	3	1.76% [1.64, 1.88]	0.24% [0.20, 0.28]	1.22% [1.12, 1.32]	0.839 [0.767, 0.911]	0.091 [0.069, 0.112]	0.570 [0.507, 0.632]
2	4	0.24% [0.20, 0.28]	0.02% [0.01, 0.04]	0.15% [0.12, 0.18]	0.088 [0.068, 0.108]	0.005 [0.001, 0.009]	0.052 [0.037, 0.067]

Table A2. Delay tail checks (R = 1 and 2; B = 2, 3, 4)

R	B (days)	Delay P95 baseline (h)	Delay P95 safety (h)	Delay P95 tail (h)	P(delay>24h) baseline	P(delay>24h) safety	P(delay>24h) tail
1	2	26.1	26.0	23.0	6.08%	5.99%	4.57%
1	3	26.1	26.0	23.0	6.08%	5.99%	4.57%
1	4	26.1	26.0	23.0	6.08%	5.99%	4.57%
2	2	32.0	32.0	29.0	9.61%	9.68%	8.27%
2	3	32.0	32.0	29.0	9.61%	9.68%	8.27%
2	4	32.0	32.0	29.0	9.61%	9.68%	8.27%