Incorporating Spatio-Temporal Weather Risk into Supply Chain Design

by

Insu Kang

Submitted on May 8, 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

Supply chains operating across global networks are increasingly exposed to recurring weather disruptions that inflate transportation lead times in predictable spatial and temporal patterns; however, most fulfillment frameworks treat these disruptions as random events and fail to leverage their seasonal recurrence for proactive planning. This capstone develops a weather-aware optimization framework that quantifies recurring weather risk and embeds it into supply chain fulfillment decisions. Using climate data and U.S. airport performance records, weather disruption probabilities are modeled and translated into lead time variability estimates across truck and air segments, as well as terminal nodes. A Mixed-Integer Linear Programming model compares a risk-agnostic Baseline policy with a Risk-Aware policy that jointly minimizes expected lead time and its standard deviation, evaluated through Monte Carlo simulation over 22,125 orders. The results show that the Risk-Aware policy reduces late orders by 63.8%—from 127 to 46—providing quantitative evidence that incorporating spatio-temporal weather risk into fulfillment decisions meaningfully improves service reliability. The analysis further reveals a seasonal concentration of disruption risk, with January and February accounting for a disproportionate share of late orders, suggesting that targeted Risk-Aware fulfillment during high-risk months can yield substantial reliability gains without incurring unnecessary lead time overhead year-round.

Capstone Advisor: Sarah Katharina Schaumann
Title: Postdoctoral Associate, Center for Transportation and Logistics

Capstone Co-Advisor: Selene Silvestri
Title: Research Scientist, Center for Transportation and Logistics

TABLE OF CONTENTS

Incorporating Spatio-Temporal Weather Risk into Supply Chain Design

1. Introduction

1.1. Background

In today's interconnected global economy, supply chains span multiple regions, transportation modes, and geopolitical environments—making them increasingly vulnerable to diverse and overlapping disruptions. Events such as extreme weather, trade policy shifts, pandemics, and port congestion have revealed the fragility of networks reliant on synchronized flows of materials, information, and capital.

Recent global crises have shown how disruptions can cascade across the system, leading to shortages, extended lead times, and escalating costs. For instance, Ericsson lost 400 million Euros after its supplier's semiconductor plant caught on fire in 2000, and Apple lost many customer orders during a supply shortage of DRAM chips after an earthquake hit Taiwan in 1999. Hendricks and Singhal (2005) report that companies suffering from supply chain disruptions experienced 33–40% lower stock returns relative to their industry benchmarks. Consequently, strengthening supply chain resilience—defined as the capacity for an enterprise to survive, adapt, and grow in the face of turbulent change—has become a strategic priority for many organizations (Fiksel, 2006).

Against this backdrop, our project focuses on developing a risk-informed optimization framework that incorporates spatio-temporal disruption patterns into supply chain routing decisions, ultimately enhancing both resilience and on-time delivery performance.

1.2. Motivation

There are two complementary approaches to building supply chain resilience: reactive strategies, which focus on mitigating disruptions after they occur, and proactive strategies, which aim to anticipate and prevent them (Grötsch et al., 2013). However, many proactive approaches still rely on static assumptions that fail to account for how certain risks recur in predictable spatial and temporal patterns (Wieland, 2021). For instance, assigning fixed disruption probabilities or permanently excluding high-risk routes can oversimplify the operational reality and prevent firms from leveraging historical recurrence patterns for more effective proactive planning.

In this context, this capstone project explores how recurring weather can be systematically incorporated into fulfillment policy. By quantifying weather-induced lead time variability using historical climate data and embedding this uncertainty into an optimization framework, the project seeks to evaluate whether a risk-informed fulfillment policy can meaningfully improve on-time delivery performance compared to a static, risk-agnostic baseline.

1.3. Problem Statement and Key Questions

Traditional risk assessments often rely on static assumptions that overlook how certain disruptions recur in predictable spatial and temporal patterns such as seasonal weather anomalies or recurring snowfall along key transportation corridors. Although many of these disruptions follow recognizable cycles, firms frequently treat them as random or isolated events. This prevents them from leveraging historical recurrence patterns to anticipate and prepare for future disruptions which in turn results in inefficient fulfillment policy.

As a result, companies struggle to translate proactive intent into timely and cost-effective operational actions. What they need instead is a dynamic framework that continuously references spatio-temporal risk patterns and adjusts decisions—such as route selection or supplier allocation—based on when and where risks are likely to reappear.

Therefore, the central question of this capstone is:

How can supply chains enhance resilience by identifying and responding to recurring disruptions across spatial and temporal dimensions?

To address this question, the study undertakes the following steps:

1. Model recurring weather disruption probabilities using historical climate data and airline performance data.
2. Translate weather disruption probabilities into lead time inflation estimates across truck and air transportation segments.
3. Implement a risk-informed fulfillment policy using Mixed-Integer Linear Programming (MILP) that incorporates both expected lead time and its variability.
4. Compare the risk-informed policy against a static baseline through Monte Carlo Simulation (MCS).
5. Evaluate both policies using key metrics such as average lead time and late order count.

1.4. Project Goals and Expected Outcomes

This project aims to develop a data-driven, weather-aware optimization framework for proactive supply chain fulfillment policy. By quantifying how recurring weather disruption events inflate transportation lead times, the framework enables risk-informed fulfillment decisions that go beyond static, risk-agnostic planning—improving both service reliability and lead time efficiency.

This project produces three main deliverables: (1) a probabilistic weather disruption model built on historical climate data and airline performance records, capturing monthly lead time inflation across both truck and air transportation; (2) an optimization model that compares a risk-agnostic Baseline policy against a Risk-Aware policy jointly minimizing expected lead time and its standard deviation; (3) and a simulation-based evaluation that compares dynamic and static fulfillment strategies in terms of service level, cost efficiency, and resilience.

The expected outcomes are threefold: (1) quantitative evidence that incorporating weather-induced lead time variability into fulfillment policy reduces late orders without a disproportionate increase in average lead time; (2) practical insights for industry managers on when and how to use spatio-temporal risk-informed decisions in fulfillment planning; (3) and a replicable, data-driven methodology that can be adapted by companies operating global networks facing recurring disruptions.

2. State of the Practice

This project aims to identify recurrent disruptions within a given global supply chain network, quantify their likelihood and severity, characterize their spatio-temporal patterns, and assess policy effectiveness using simulation. To establish the theoretical foundation, we review the following five domains: (1) supply chain network design, (2) supply chain risk identification and classification, (3) risk quantification, (4) risk mitigation strategies, and (5) simulation methods for policy evaluation.

2.1. Supply Chain Network Design

A supply chain network consists of suppliers, fulfillment centers or distribution centers, and customers connected by transportation lanes. Supply chain network design (SCND) determines this structure through facility location, capacity allocation, sourcing relationships, and lane configuration. It is a central determinant of cost and service performance (Melo et al., 2009), as well as of disruption exposure. Properties such as node redundancy and lane diversity shape both the likelihood of propagation and the speed of recovery (Snyder et al., 2016). Substantial literature accordingly embeds disruption uncertainty directly into network design through robust and stochastic optimization (Govindan et al., 2017).

2.2. Supply Chain Risk Identification and Classification

Identifying risks is the essential starting point for any supply chain risk management process. Risk identification should account for the broader context in which the supply chain operates, because risk manifestation is strongly shaped by contextual factors such as region and industry (Schleper et al., 2024). Prior research outlines several methodologies for this purpose, which generally fall into four groups: (1) listing-based methods, where past disruption events are reviewed to infer potential future risks; (2) taxonomy-oriented approaches, which establish a structured framework to systematically capture and categorize risks; (3) scenario analyses, which evaluate critical risk drivers and their implications for supply chain performance to support the development of operational contingency plans; and (4) risk mapping techniques, which visualize potential points of vulnerability within the supply chain before disruptions occur (Singhal et al., 2011).

Structured risk classification plays a critical role in translating identified risks into actionable mitigation strategies (McKinsey & Company, 2020). The supply chain literature generally organizes risks into a few broad categories, most commonly distinguishing between (1) external risks—those arising from environmental, geopolitical, and market forces—and (2) internal risks stemming from a firm's own processes and operations (Jüttner, 2005). Among external risks, natural hazards such as extreme weather events are particularly relevant because they tend to recur in predictable spatial and temporal patterns (Tang, 2006). Unlike one-off geopolitical or market disruptions, weather-related risks follow seasonal cycles that can be modeled using historical data—a property this project directly exploits to build probabilistic delay estimates for supply chain fulfillment policy.

2.3. Risk Quantification

In order to assess and compare alternative solutions that reduce risk impact, decision makers must be able to quantify risk. Following the widely accepted definition of risk as the product of likelihood and impact (Manuj & Mentzer, 2008), this review distinguishes two broad categories of approaches used to measure risk: (1) qualitative/semi-quantitative and (2) quantitative (Rinaldi et al., 2022). In this review, we focus on quantifying event-level likelihood and severity as intrinsic properties of disruption risks.

Qualitative methods (such as expert elicitation, Delphi, and ordinal severity scales) are useful when data is sparse, though they depend heavily on subjective judgment. Quantitative methods rely on statistically grounded models that use historical or empirical data to characterize disruption magnitude with greater precision. The most common technique is statistical distribution fitting, which models delay increments, capacity loss, recovery time, or downtime using Gamma, Weibull, Lognormal, or empirical (non-parametric) distributions. Bayesian inference methods are also used in settings where disruption data are sparse, allowing prior knowledge and observed evidence to be combined to estimate the likelihood or severity parameters more robustly (Hosseini & Ivanov, 2020). Together, these approaches provide a structured foundation for quantifying the inherent likelihood and severity of disruption events prior to assessing their downstream impact on the broader supply chain.

2.4. Risk Mitigation Strategies

The literature generally classifies risk-mitigation strategies into three broad categories: (1) avoidance, which eliminates exposure to risky nodes, suppliers, or processes; (2) mitigation, which reduces the likelihood or severity of disruptions through measures such as supplier diversification or infrastructure reinforcement; and (3) absorption or resilience, which improves a system's ability to buffer or adapt to shocks with tools like safety stock or flexible capacity

(Hajmohammad & Vachon, 2015; Sheffi & Rice, 2005). These categories are well established in supply chain risk research and represent the foundational mechanisms through which firms reduce vulnerability.

These strategies can be applied at strategic, tactical, and operational levels (Christopher & Peck, 2004), with operational responses such as re-routing being particularly relevant to this project (Ivanov et al., 2016). The Risk-Aware policy developed in this study falls within the mitigation category at the operational level: it reduces the impact of weather disruptions by proactively selecting lower-variability routes rather than eliminating or absorbing risk after the fact.

2.5. Simulation Methods for Evaluating Supply Chain Operational Policies

Because supply chain disruptions propagate through networks in nonlinear and time-dependent ways, analytical or steady-state approximations often fail to capture their true operational impact. Delays at one node can trigger congestion, queue buildup, capacity imbalances, and cascading downstream effects—phenomena widely documented in ripple-effect studies of supply chain networks (Ivanov & Dolgui, 2020; Gao et al., 2019). As a result, simulation has become a central methodological tool in the supply chain risk literature, enabling researchers to translate event-level disruption severity into system-level performance outcomes while accounting for uncertainty, temporal dynamics, and network interactions.

A wide range of simulation approaches has been employed to study disruption propagation and resilience. MCS repeatedly samples disruption events to produce probabilistic impact distributions and tail-risk metrics such as Conditional Value at Risk (CVaR) (Schmitt & Singh, 2009; Heidary & Aghaie, 2015). In practice, Monte Carlo methods are often layered on top of Discrete-Event Simulation (DES) or network-based optimization models. The outer loop samples disruption scenarios while the inner loop evaluates their operational impact.

While System Dynamics and Agent-Based Simulation are also common in supply chain research, MCS is most natural for this project given the Bernoulli-based disruption model. This disruption model relies on independent Bernoulli trials for weather events and airline delays, making repeated probabilistic sampling the most natural evaluation method.

2.6. Summary of Findings and Relevance to the Project

The literature on supply chain network design highlights that network structure—facility location, lane configuration, and node redundancy—is a fundamental determinant of both efficiency and disruption exposure. Complementing this perspective, the supply chain risk literature emphasizes the importance of structured risk identification, estimation of disruption likelihood and severity, and the evaluation of mitigation strategies. Prior work also demonstrates that disruptions can propagate through networks in nonlinear ways and that simulation is frequently used to analyze such system-level dynamics.

Building on these insights, this project adopts the supply chain network decision structure as the underlying operational context but employs a simplified modeling scope that excludes propagation effects and dynamic interactions. Instead, the focus is on probabilistic characterization of recurring disruptions and the use of simulation to generate uncertain delay realizations for comparing baseline and risk-informed fulfillment policies.

3. Methodology

This study develops a framework to evaluate the operational impact of recurring weather disruptions on supply chain network design decisions. The framework consists of three methodological components: (1) probabilistic delay modeling, (2) MILP-based optimization, and (3) MCS-based performance evaluation. It compares a Baseline policy that ignores disruption risk with a Risk-Aware policy that incorporates disruption-induced delays.

First, weather disruptions are modeled using the climate and airport performance data. Each disruption is defined as a probabilistic event that inflates transportation lead time and is modeled separately across two layers: truck segments (arcs) and U.S. hub airports (nodes).

Second, lead time optimization is performed using MILP. The Baseline policy minimizes nominal lead time without incorporating weather delays, while the Risk-Aware policy additionally penalizes lead time variability by incorporating a standard deviation term into the objective function. The Risk-Aware policy solves a separate MILP for each of the 12 months to capture month-specific weather and airline delay patterns.

Third, MCS is conducted on the routes determined by the MILP to evaluate actual on-time delivery performance. For each order, weather event occurrences and airline delays are sampled probabilistically to compute the realized lead time, which is then compared against the promised delivery window to assess the performance of both policies.

3.1. Optimization Model

Supply path selection is performed using MILP. The decision variable $x[k, st, s, h, d]$ is a binary variable that takes the value 1 if the combination of customer cluster (k), storage type (st), supplier (s), U.S. hub airport (h), and distribution center (d) is selected, and 0 if it is not.

3.1.1. Baseline Policy

The Baseline policy assigns each (cluster, storage-type) demand to the supply path with the shortest nominal lead time, ignoring weather disruptions. It represents a conventional, disruption-agnostic decision rule in which the optimizer relies solely on the nominal lead time reported in the network data. The objective is:

$$\min \sum_{k,st,s,h,d} \mathrm{LT} \cdot x[k, st, s, h, d] \qquad (1)$$

This is equivalent to setting all delay probabilities to zero in the stochastic formulation that is introduced next.

3.1.2. Risk-Aware Policy

The Risk-Aware policy replaces the nominal lead time with the expected lead time under the stochastic delay model, and augments the objective with a standard deviation penalty. Paths that are fast on average but exhibit high variability are therefore discouraged in favor of routes that are both fast and reliable:

$$\min \sum_{k,st,s,h,d} (E[LT] + \lambda \cdot \sigma[LT]) \cdot x[k, st, s, h, d] \qquad (2)$$

where $E[LT]$ is the expected lead time incorporating both airline and weather expected delays, and $\sigma[LT]$ is computed analytically using the sum of variance of independent Bernoulli random variables. Since each weather event and airline delay is modeled as an independent Bernoulli trial, the variance is expressed as:

$$\sigma^2[LT] = \sum_i p_i\,(1 - p_i) \cdot f_i^2 \qquad (3)$$

where p_i is the probability of event i and f_i is the corresponding slowdown factor or delay severity. λ is the risk aversion coefficient set to a default value of 2.0; larger values induce stronger preference for routes with lower lead time variability.

The Risk-Aware policy solves a separate MILP for each of the 12 months, deriving month-specific optimal routes that reflect seasonal weather patterns and airline delay characteristics.

3.1.3. Key Constraints

Each (cluster, storage type) pair must be assigned to exactly one route:

$$\sum_{s,h,d} x[k, st, s, h, d] = 1, \quad \forall k, st \qquad (4)$$

The total demand $D_{k,st}$ assigned to each DC must not exceed its monthly throughput capacity by storage type:

$$\sum_{\mathrm{k,s,h}} \mathrm{D}_{k,st} \cdot x[k, st, s, h, d] \leq C_{\mathrm{d,st}}^{\mathrm{DC}}, \quad \forall d, st \qquad (5)$$

The total demand assigned to each supplier must not exceed its monthly supply capacity by storage type:

$$\sum_{\mathrm{k,h,d}} \mathrm{D}_{k,st} \cdot x[k, st, s, h, d] \leq C_{\mathrm{s,st}}^{SUP}, \quad \forall s, st \qquad (6)$$

Finally, to improve computational efficiency, only the top 200 lowest-cost route candidates per (cluster, storage type) combination are included as MILP decision variables. This significantly reduces the number of variables while minimizing loss of solution optimality in practice.

3.2. Delay Modeling

For brevity, the following segment abbreviations are used throughout: Supplier-to-Terminal (S2T), Terminal-to-Terminal (T2T), Terminal-to-DC (T2D), DC-to-Customer (D2C), and Supplier-to-DC (S2D).

3.2.1. Weather Delay: Truck Segments (Arcs)

Weather-induced delays on truck segments are defined across four event categories based on operational thresholds that reflect the conditions under which truck transportation is meaningfully affected.

Rain thresholds follow Wang et al. (2021), who define moderate rain as ≥ 5 mm/day and heavy rain as ≥ 20 mm/day. Snow thresholds follow the Local Winter Storm Scale (Cerruti & Decker, 2011): moderate snow at ≥ 5.1 mm SWE/day and significant snow at ≥ 10.2 mm SWE/day, converted from physical depth using a standard 10:1 snow-to-liquid ratio. Speed reduction rates are adopted from the Federal Highway Administration (FHWA) Road Weather Management Program, using the upper bound of the Average Speed column: 13% for moderate rain and moderate snow, 16% for heavy rain, and 40% for significant snow, as summarized in Table 1.

Event	Variable	Threshold	Speed Reduction
Moderate rain	Precipitation	≥ 5mm/day	-13%
Heavy rain	Precipitation	≥ 20mm/day	-16%
Moderate snow	Snowfall	≥ 5.1mm SWE/day	-13%
Significant snow	Snowfall	≥ 10.2mm SWE/day	-40%

Table 1. Weather Variables With the Threshold and the Impact

The slowdown factor f is derived from speed reduction rate r as follows:

$$f = \frac{1}{1 - r} - 1 \qquad (7)$$

This factor represents the fractional increase in lead time required to traverse the same distance under adverse weather conditions. The slowdown factors for each event are as shown in Table 2.

Event	Speed Reduction	Slowdown Factor
Moderate rain	−13%	0.149
Heavy rain	−16%	0.190
Moderate snow	−13%	0.149
Significant snow	−40%	0.667

Table 2. Weather Variables With the Impact and Slowdown factor

Rather than averaging delay values only at the origin and destination, this study applies a Route Corridor approach in which intermediate points are sampled every 100 miles along the great circle path between two nodes, and delay values are averaged across all sampled points. This captures the weather conditions encountered along the full length of long-haul truck routes, such as those crossing the continental United States. The Route Corridor approach is applied to T2D, D2C, and S2D segments.

The monthly weather delay probability for each segment is computed by dividing the number of event days in a given month by the total number of days in that month:

$$p_{event}(m, arc) = \frac{N_{event}(m, arc)}{D_m} \qquad (8)$$

where $N_{event}(m, arc)$ is the number of days in month m during which the weather event occurs along the segment, and D_m is the total number of days in month m.

3.2.2. Hours-of-Service (HOS) Regulatory Adjustment

Under U.S. Federal Motor Carrier Safety Administration regulations, truck drivers are required to take a mandatory 10-hour rest period after every 11 hours of continuous driving. Accordingly, lead times are adjusted from pure driving time to legally compliant calendar time. For example, a route requiring 21 hours of driving incurs one mandatory rest period, resulting in an actual elapsed time of 31 hours.

3.2.3. Air Transportation Delay: Hub Airport Nodes

Since direct observation of node-level ground delays and bottlenecks is not feasible, delays at U.S. hub airport nodes are modeled using the weather-induced cancellation rate as proxy. Airports with higher cancellation rates are assumed to be more prone to ground delays, gate congestion, and processing bottlenecks, which are estimated as a fixed penalty of 48 hours in lead time.

The weather-induced arrival delay of a flight is modeled using a separate Bernoulli trial with the monthly weather delay probability. If triggered, the corresponding average delay severity is added to the lead time.

3.3. Simulation and Performance Evaluation

MCS is conducted on the routes determined by the MILP to evaluate realized on-time delivery performance. Unlike the expectation-based route selection in MILP, the simulation explicitly samples probabilistic delays at the individual order level.

3.3.1. Simulation Structure

For each order, the realized lead time is computed in the following sequence:

1. T2D segment (hub airport to DC by truck): An independent Bernoulli trial is conducted for each of the four weather events (moderate rain, heavy rain, moderate snow, significant snow). The applicable slowdown factors are cumulatively applied to the base T2D lead time, followed by an HOS adjustment.
2. Hub airport node (airline delay): Two independent Bernoulli trials are conducted. The first, based on the cancellation rate, determines whether a node congestion penalty (48 hours) is applied. The second, based on the weather delay probability, determines whether the average weather delay severity is added to the lead time.
3. D2C segment (DC to customer by truck): Weather events are applied using the same procedure as the T2D segment.
4. S2D segment (U.S. supplier to DC by truck): For domestic routes that bypass hub airports, the same procedure as in step 1 is applied to the S2D lead time in place of T2D. No airline delay is incurred, since the hub airport is skipped.

The total realized lead time is computed as:

$$LT_{eff}^{global} = LT^{S2T+T2T} + LT_{airport}^{node} + LT^{T2D} + LT^{D2C} \qquad (9)$$

$$LT_{eff}^{domestic} = LT^{S2D} + LT^{D2C} \qquad (10)$$

3.3.2. Performance Metrics

An order is classified as late if its realized lead time exceeds the promised delivery window (0.215 weeks, approximately 36 hours). The two policies are compared on two metrics:

- Average Realized Lead Time: The mean realized lead time across all orders.
- Late Order Count: The absolute number of orders exceeding the promised lead time.

3.4. Sensitivity Analysis Design

Sensitivity analysis is conducted to assess the robustness of the results and to identify the conditions under which the Risk-Aware policy provides the greatest operational benefit. The analysis focuses on two key parameters:

- Risk aversion coefficient (λ) (default: 2.0, range: 1.0–3.0): Controls the degree to which lead time variability is penalized in the MILP objective. Higher values induce stronger preference for stable routes at the cost of higher average lead time.
- Node congestion penalty (α) (default: 48 hours, range: 24–72 hours): Governs how severely flight cancellation rates translate into lead time inflation at hub airport nodes.

4. Result and Discussion

4.1. Data and Inputs

4.1.1. Climate Data

Weather data are obtained from the European Centre for Medium-Range Weather Forecasts Reanalysis 5th generation (ERA5), post-processed daily statistics on single levels from 1940 to present. ERA5 provides global daily weather observations at a native resolution of 0.25-degree grid, and this study uses two variables: total precipitation (mm/day) and snowfall (mm snow water equivalent/day). Daily observations are aggregated into monthly event-day counts. For each grid cell, daily observations are compared against operational thresholds to produce binary event flags. These flags are then aggregated by month to yield the number of event days per grid cell, per month.

4.1.2. Airport Performance Data

The Federal Aviation Administration's Airline Service Quality Performance System (ASQP) Causal Report as well as the Bureau of Transportation Statistics' (BTS) airport performance data are combined to derive monthly arc and node delay metrics for the 22 U.S. hub airports in the network.

The weather delay probability is computed as follows. The share of weather-caused delays among all delayed operations is obtained from BTS on a per-airport, per-month basis for all reporting carriers. This share is then multiplied by the number of delayed flights reported in ASQP and divided by the total number of arriving flights to derive the monthly delay probability. Since ASQP covers only large carriers that are subject to mandatory reporting, combining it with the BTS weather delay share corrects for coverage limitations.

The average weather delay severity is computed from ASQP data as the mean weather-caused delay duration per affected flight, aggregated by airport and month.

The flight cancellation rate is derived from ASQP data and is used as a proxy for estimating additional operational delays at the airport node—including ground delays, gate congestion, and processing bottlenecks—rather than as a direct measure of flight cancellations. These three metrics (weather delay probability, average delay severity, and cancellation rate) are aggregated on an arriving-flight basis and are used to model airport-level delays at each U.S. hub.

4.1.3. Supply Chain Network and Order Data

The supply chain network being studied consists of 20 suppliers, 22 U.S. hub airports, 40 distribution centers (DCs), and 125 customer clusters. Suppliers are categorized by storage type (Rack, Flammable, Temp and Secured). Two route types exist: global routes (overseas supplier-hub airport-DC-customer) and domestic routes (U.S. supplier-DC-customer). Each node carries geographic coordinates, enabling integration with ERA5 climate data.

Historical orders used for policy evaluation comprise 22,125 records aggregated to the cluster level and grouped by storage type. Each order carries a monthly timestamp, a customer cluster ID, and a storage type, allowing monthly MILPs and simulations to reference month-specific demand patterns.

4.2. Overall Policy Comparison

Table 3 summarizes the performance of the Baseline and Risk-Aware policies across the two key metrics evaluated over 22,125 orders, with Figure 1 providing a visual comparison of late orders and average lead time.

Policy	Avg Lead Time	Late Orders
Baseline	9.55 hours	127
Risk-Aware (λ =2.0)	10.21 hours	46

Table 3. Overall Performance Comparison: Baseline vs. Risk-Aware

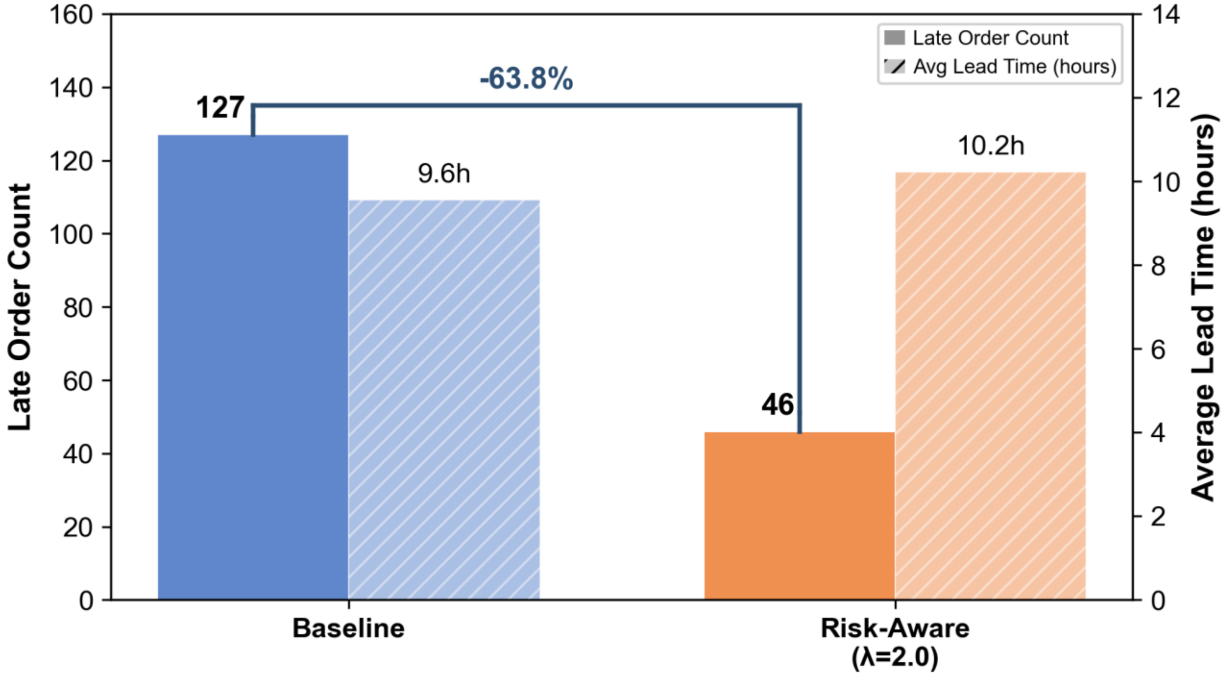

Figure 1. Late Order Count and Average Lead Time comparison

The Risk-Aware policy reduces late orders from 127 to 46, a 63.8% reduction. This improvement comes at the cost of a modest increase in average lead time of approximately 0.7 hours.

The increase in average lead time under the Risk-Aware policy is a structural consequence of the optimization objective. By penalizing lead time variability through the term $\lambda \cdot \sigma[LT]$, the Risk-Aware policy systematically favors routes with lower uncertainty—even when such routes carry a nominally longer expected lead time. In other words, the policy trades a small, predictable increase in average lead time for a meaningful reduction in the tail risk of late deliveries. This trade-off reflects the core value proposition of Risk-Aware fulfillment policy: sacrificing marginal efficiency for improved reliability.

4.3. Lead Time Distribution Analysis

Beyond average lead time, the two policies differ meaningfully in their lead time distributions, as illustrated in Figure 2. The Risk-Aware policy exhibits a lower standard deviation (4.29 hours vs. 4.89 hours for Baseline), confirming that the variability penalty in the objective function translates into more stable realized lead times across orders.

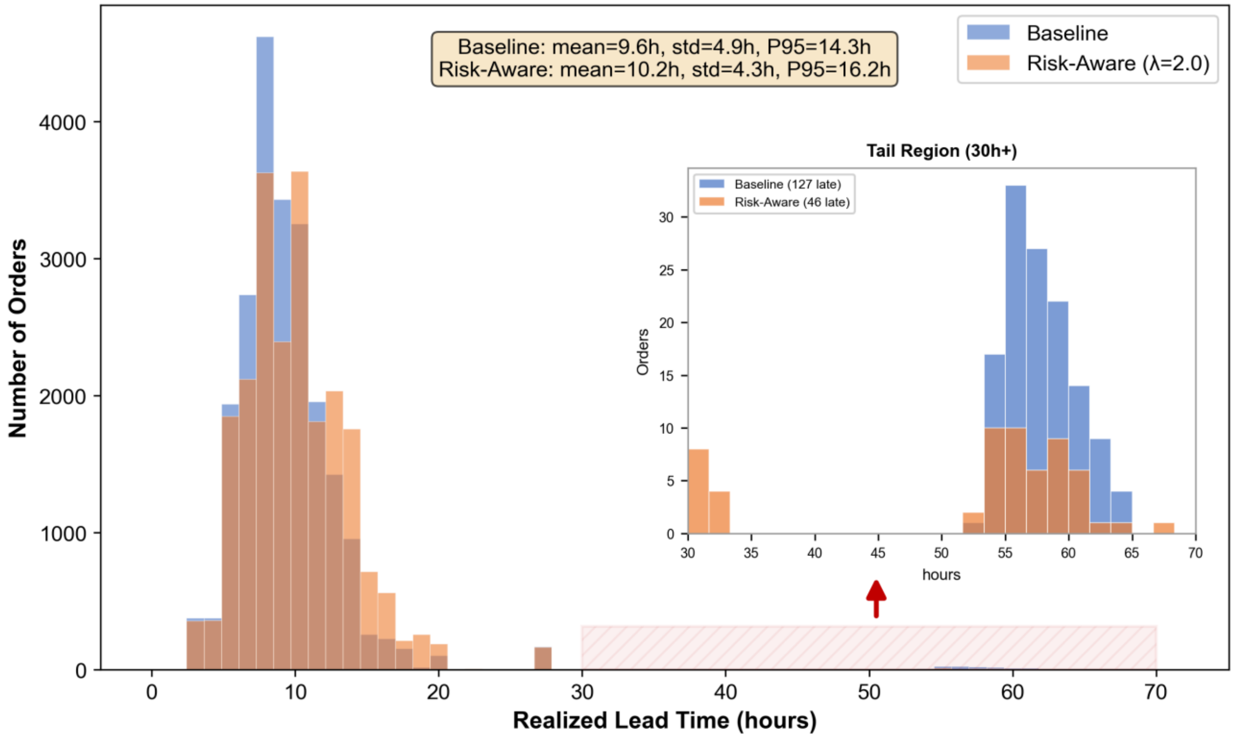

Figure 2. Realized Lead Time Distribution Comparison

At the 95th percentile, the Baseline policy yields a lead time of approximately 14.3 hours, compared to 16.2 hours under the Risk-Aware policy. This higher 95th percentile value does not indicate worse tail performance; rather, it reflects the fact that Risk-Aware routes are systematically longer on average due to their preference for lower-variability corridors. Critically, the Risk-Aware policy produces far fewer orders that exceed the promised delivery window (46 vs. 127), confirming that the true tail risk—the frequency of late deliveries—is substantially reduced.

The key distinction is not the magnitude of extreme outcomes, but their frequency. Risk-Aware policies substantially reduce the likelihood of such disruptions, resulting in a more stable and reliable system.

4.4. Seasonal Pattern Analysis

Late orders under the Baseline policy are heavily concentrated in the winter months, as shown in Figure 3, with January and February accounting for 24 and 18 late orders respectively. This pattern is consistent with the higher frequency of snow-related weather events during these months, which inflate truck segment lead times along routes that the Baseline policy selects without accounting for seasonal risk.

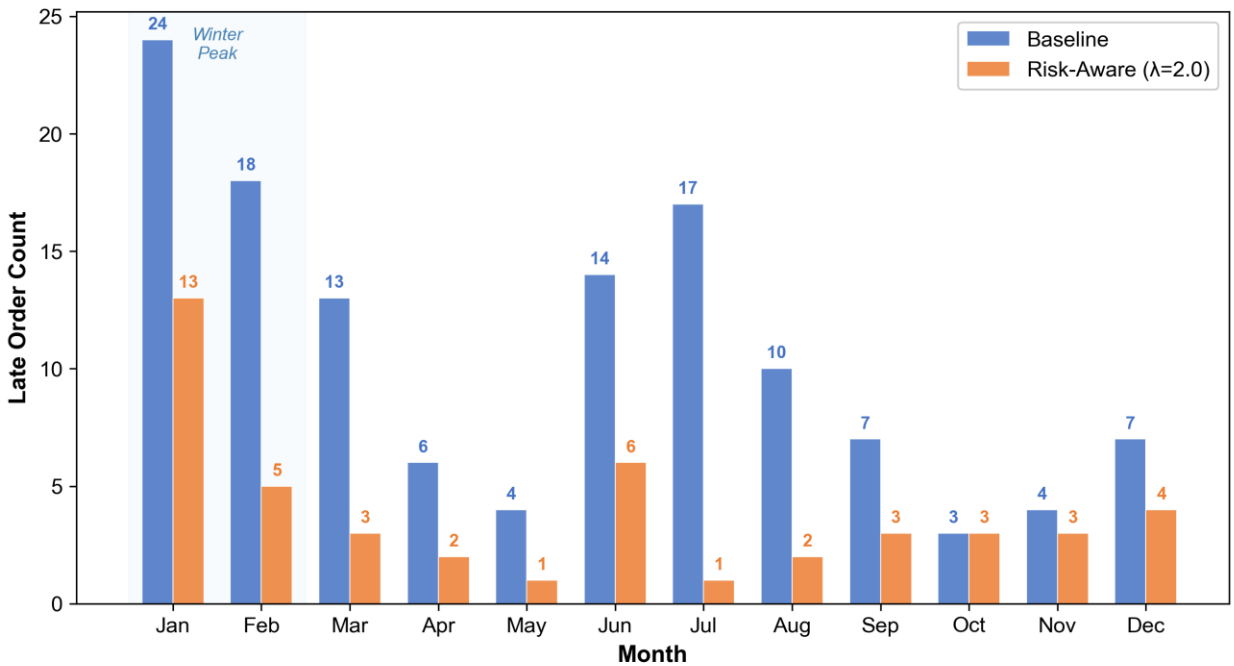

Figure 3. Monthly Late Order Count Comparison

The Risk-Aware policy reduces late orders in January from 24 to 13 and reduces July delays from 17 to 1. This improvement is driven by month-specific optimization, which incorporates seasonal weather risks and airline delay patterns to proactively reroute shipments away from high-risk corridors. The remaining late orders are less concentrated, indicating that residual delays are primarily driven by stochastic variability rather than structural inefficiencies in fulfillment decisions.

4.5. Route Switching Analysis

Of the 500 (cluster, storage type) pairs in the network, 250 (50.0%) are assigned to different routes under the Risk-Aware policy compared to the Baseline. This substantial degree of route switching, illustrated in Figure 4, reflects the meaningful differences in how the two policies evaluate route cost.

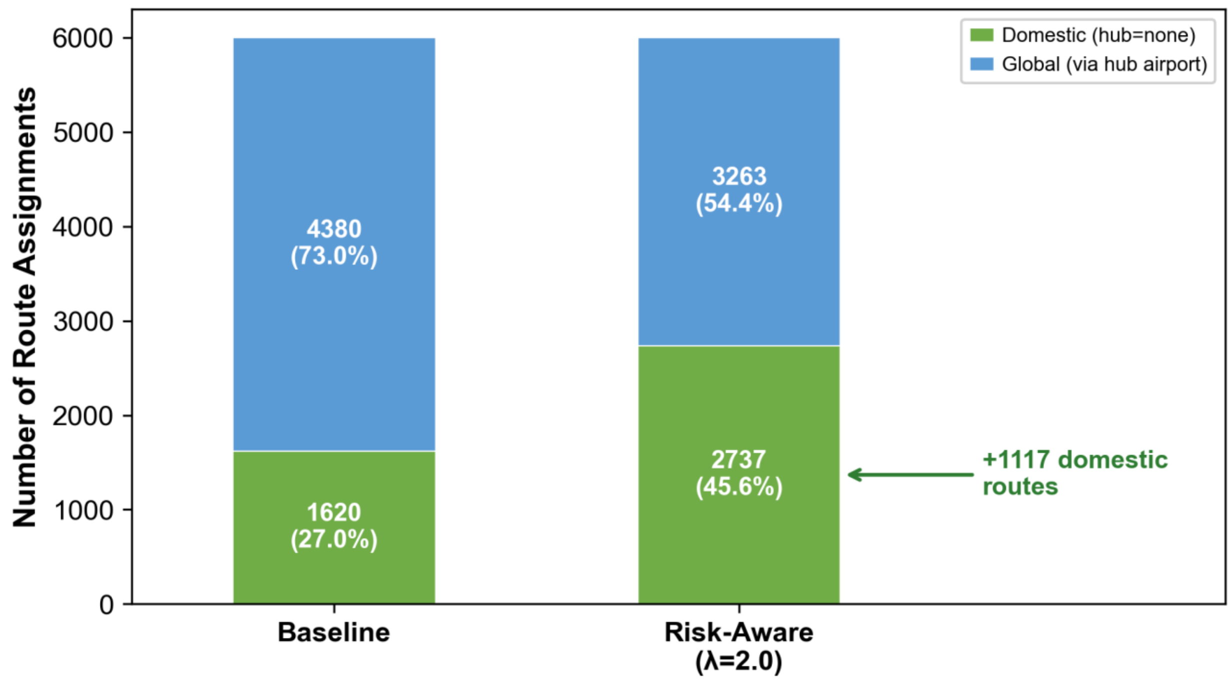

Figure 4. Domestic vs. Global Route Assignment Share Showing Shift Toward Domestic Routes

A notable shift is observed in the use of domestic routes (i.e., those bypassing hub airports), which increase from 1,620 to 2,737 route assignments under the Risk-Aware policy. This suggests that the Risk-Aware policy systematically avoids U.S. hub airports when domestic alternatives are available, as hub airports introduce airport-level delay risk that is explicitly penalized in the objective function.

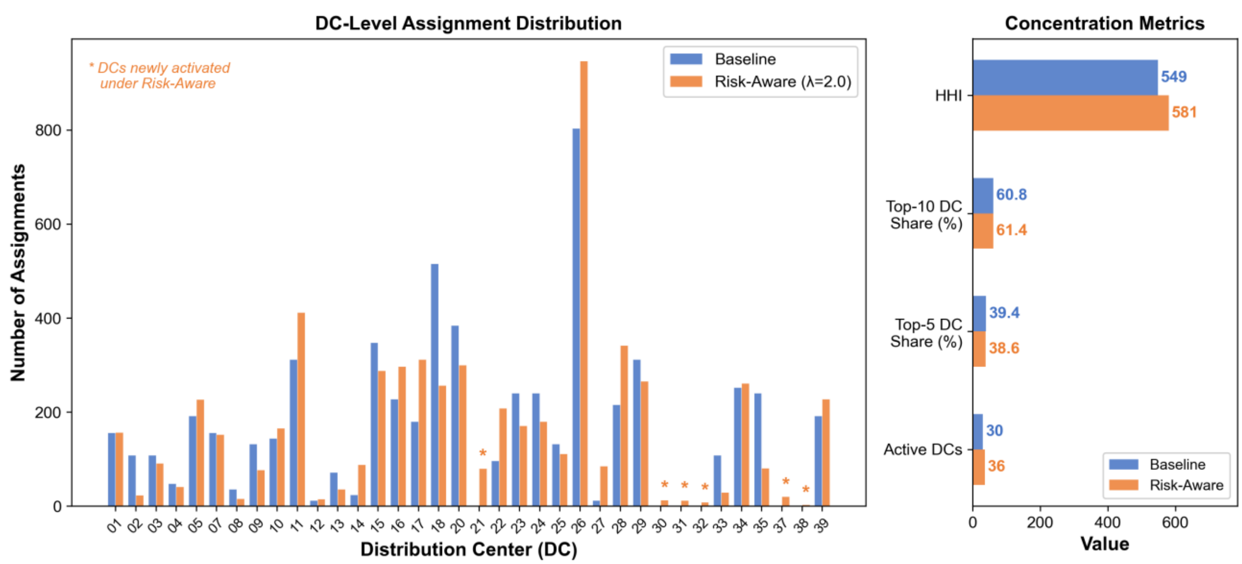

Figure 5. DC Assignment Pattern

Figure 5 illustrates the DC assignment pattern under the two policies. The number of active DCs increases from 30 to 36, with six previously unused DCs newly activated. Meanwhile, the Herfindahl-Hirschman Index (HHI) of assignment shares—a concentration measure where higher values mean demand is routed through fewer DCs—rises modestly from 549 to 581, and the top-five DC share declines only slightly (39.4% to 38.6%). In other words, the Risk-Aware policy activates more DCs overall, but the additional volume tends to cluster on a smaller subset of low-weather-risk corridors rather than distributing evenly. This pattern is an emergent outcome of individual route-level optimization: when every leg—including inbound S2D—is variance-penalized, certain DCs in stable corridors become disproportionately attractive, and the solver concentrates assignments onto them whenever capacity permits.

4.6. Sensitivity Analysis Results

Sensitivity analysis is performed on the risk-aversion coefficient (λ) and the cancellation penalty (α), varying each in turn while holding the other at its default. Results are summarized in Figure 6.

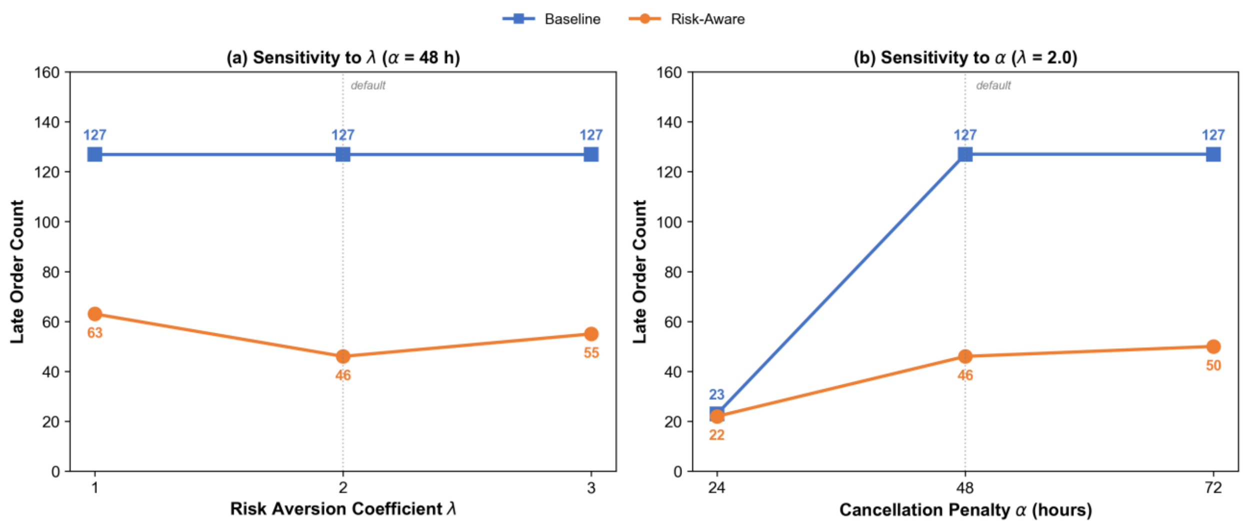

Figure 6. Sensitivity Analysis

As shown in Figure 6a, with α fixed at 48 hours, Risk-Aware late orders fall from 63 at λ = 1 to 46 at λ = 2, then rise to 55 at λ = 3, while the mean realized lead time climbs at every step (9.74, 10.21, and 10.62 hours) and the standard deviation moves from 4.28 to 4.29 to 4.63 hours. The two transitions are qualitatively different. From λ = 1 to λ = 2, the policy swaps hub-routed paths—whose variance is dominated by the 48-hour cancellation penalty—for domestic paths that bypass the hub. This is a tail-cutting substitution that removes a rare 48-hour shock, so the standard deviation barely moves yet lateness drops sharply. From λ = 2 to λ = 3, the high-risk hub routes are exhausted, and further variance reduction is achieved only by pushing already-domestic clusters onto longer truck lanes. This variance-chasing substitution pulls the mean toward the 36-hour deadline and inflates the standard deviation in absolute terms, so lateness rebounds. Therefore, λ = 2 sits near the operational sweet spot. The Baseline is invariant to λ and remains at 127.

The α sweep in Figure 6b shows a different pattern. With λ fixed at 2, the two policies are nearly identical at α = 24 hours (Baseline 23, Risk-Aware 22) because the mild node penalty rarely pushes realized lead time past the 36-hour window. As α grows to 48 hours the Baseline jumps to 127 while Risk-Aware climbs only to 46 (−63.8%), and at α = 72 hours the gap is preserved (127 vs. 50, −60.6%). The Baseline's realized standard deviation inflates from 3.73 to 6.36 hours across this range, whereas Risk-Aware contains the growth (3.71 to 5.27 hours) by shifting demand to domestic routes (2,071, 2,737, 3,211 at α = 24, 48, 72 h).

Overall, The Risk-Aware advantage is conditional on disruption severity: when hub cancellations are cheap both policies are similar, but as α grows Risk-Aware gains by tail-cutting away from hubs. The λ sweep shows this benefit saturates—beyond $\lambda \approx 2$, the policy starts selecting routes whose average lead time approaches the promised window, so lateness is no longer caused by rare cancellation events but by the average lead time itself.

5. Recommendations

5.1. Operational Implication

The results provide clear quantitative support for adopting the Risk-Aware policy in operations. A 63.8% reduction in late orders with an average lead time increase of only 0.7 hours represents a favorable trade-off in most fulfillment contexts, where the cost of a late delivery—in terms of customer dissatisfaction, expediting costs, and service level penalties—can exceed the marginal cost of a longer average lead time.

The risk aversion coefficient (λ) provides a practical operational lever that can be calibrated to reflect the firm's risk tolerance. A higher λ value induces stronger preference for route stability at the cost of higher average lead time, while a lower value moves the policy closer to the risk-agnostic Baseline. Firms operating under tight service level agreements may benefit from higher λ settings, while those prioritizing lead time efficiency may prefer lower values.

5.2. Seasonal Fulfillment Strategy

The seasonal concentration of late orders under the Baseline policy suggests that a hybrid approach may be operationally effective—applying Risk-Aware policy during high-risk winter months (particularly January and February) while operating closer to the Baseline during low-risk periods.

This would allow firms to capture the reliability benefits of risk-informed policy where they matter most, while minimizing the average lead time overhead during seasons when weather disruption risk is low.

To operationalize this, it is recommended that month-specific MILPs be re-solved on a rolling basis using updated climate and airline performance data. This ensures that fulfillment decisions continuously reflect current seasonal risk patterns rather than relying on static historical averages.

5.3. Limitations

Several limitations of the current framework should be acknowledged.

First, in the absence of direct ground delay data, the cancellation rate derived from ASQP data is used as a proxy for node-level operational congestion at hub airports. Therefore, the assumed 48-hour penalty is an approximation, and the actual relationship between cancellation rates and ground delay duration may vary across airports and seasons.

Second, there is no single authoritative study providing the precise thresholds and speed reduction rates associated with weather-induced truck delays. The parameters used in this study—including precipitation thresholds (Wang et al., 2021), snow thresholds from the Local Winter Storm Scale (Cerruti & Decker, 2011), and truck speed reduction rates (FHWA, n.d.)—are drawn from multiple authoritative sources and synthesized into a unified framework. While each individual source is peer-reviewed or operationally published, differences in measurement context introduce residual uncertainty. This is partially addressed through sensitivity analysis on the risk aversion coefficient (λ) and node congestion penalty (α).

Third, weather exposure along truck segments is computed using great circle paths derived from the Haversine formula, which represent straight-line distances between nodes. Actual truck routes follow road networks that are constrained by highway infrastructure, terrain, and regulatory restrictions, and which may differ meaningfully from great circle paths—particularly in regions with complex topography. Utilizing the Python library OSMnx to replace great circle approximations with road network-based routing would improve the accuracy of weather exposure estimates along truck corridors.

Lastly, disruption severity is modeled using average values rather than full empirical distributions. While disruption occurrence is modeled probabilistically, the magnitude of delay conditional on occurrence is represented by a fixed average value. This simplification improves tractability and enables analytical variance computation. However, it may underestimate the tail behavior associated with extreme disruption events.

5.4. Future Research

Several directions for future research emerge from this study. First, an adaptive λ adjustment mechanism could be developed to dynamically update the risk aversion coefficient based on terminals' resilience. Second, transportation cost could be incorporated as an additional term in the objective function, enabling a multi-objective optimization framework that explicitly balances lead time efficiency, delivery reliability, and logistics cost. Third, replacing the Haversine-based route corridor with actual road network routing data would improve the fidelity of weather exposure modeling along truck segments. Fourth, the current framework focuses exclusively on weather-related disruptions. Extending the model to incorporate additional risk categories—such as geopolitical disruptions, port congestion, or supplier-level failures—would broaden its applicability and improve robustness across a wider range of real-world supply chain conditions.

References

Cerruti, B. J., & Decker, S. G. (2011). The Local Winter Storm Scale: A measure of the intrinsic ability of winter storms to disrupt society. *Bulletin of the American Meteorological Society, 92*(6), 721–737. https://doi.org/10.1175/2010BAMS3191.1

Christopher, M., & Peck, H. (2004). Building the Resilient Supply Chain. *The International Journal of Logistics Management, 15*(2), 1–14. https://doi.org/10.1108/09574090410700275

Federal Highway Administration. (n.d.). *How do weather events affect roads?* U.S. Department of Transportation. https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm

Fiksel, J. (2006). Sustainability and resilience: toward a systems approach. *Sustainability: Science, Practice and Policy, 2*(2), 14–21. https://doi.org/10.1080/15487733.2006.11907980

Gao, S. Y., Simchi-Levi, D., Teo, C.-P., & Yan, Z. (2019). Disruption Risk Mitigation in Supply Chains: The Risk Exposure Index Revisited. *Operations Research, 67*(3), 831–852. https://doi.org/10.1287/opre.2018.1776

Govindan, K., Fattahi, M., & Keyvanshokooh, E. (2017). Supply Chain Network Design under uncertainty: a Comprehensive Review and Future Research Directions. *European Journal of Operational Research, 263*(1), 108–141. https://doi.org/10.1016/j.ejor.2017.04.009

Grötsch, V. M., Blome, C., & Schleper, M. C. (2013). Antecedents of proactive supply chain risk management – a contingency theory perspective. *International Journal of Production Research, 51*(10), 2842–2867. https://doi.org/10.1080/00207543.2012.746796

Hajmohammad, S., & Vachon, S. (2015). Mitigation, Avoidance, or Acceptance? Managing Supplier Sustainability Risk. *Journal of Supply Chain Management, 52*(2), 48–65. https://doi.org/10.1111/jscm.12099

Heidary, M. H., & Aghaie, A. (2015). Risk measurement in the global supply chain using Monte-Carlo simulation. *Journal of Industrial Engineering and Management Studies, 2*(2), 1–12. https://jiems.icms.ac.ir/article_15291.html

Hendricks, K. B., & Singhal, V. R. (2005). An Empirical Analysis of the Effect of Supply Chain Disruptions on Long-Run Stock Price Performance and Equity Risk of the Firm. *Production and Operations Management, 14*(1), 35–52. https://doi.org/10.1111/j.1937-5956.2005.tb00008.x

Hosseini, S., & Ivanov, D. (2020). Bayesian Networks for Supply Chain Risk, Resilience and Ripple Effect Analysis: A Literature Review. *Expert Systems with Applications, 161,* 113649. https://doi.org/10.1016/j.eswa.2020.113649

Ivanov, D., Dolgui, A., Sokolov, B., & Ivanova, M. (2016). Disruptions in supply chains and recovery policies: state-of-the art review. *IFAC-PapersOnLine, 49*(12), 1436–1441. https://doi.org/10.1016/j.ifacol.2016.07.773

Ivanov, D., & Dolgui, A. (2020). Viability of intertwined supply networks: Extending the supply chain resilience angles towards survivability. A position paper motivated by COVID-19 outbreak. *International Journal of Production Research, 58*(10), 2904–2915. https://doi.org/10.1080/00207543.2020.1750727

Jüttner, U. (2005). Supply chain risk management: understanding the business requirements from a practitioner perspective. *The International Journal of Logistics Management, 16*(1), 120–141. https://doi.org/10.1108/09574090510617385

Manuj, I., & Mentzer, J. T. (2008). Global Supply Chain Risk Management Strategies. *International Journal of Physical Distribution & Logistics Management, 38*(3), 192–223. https://doi.org/10.1108/09600030810866986

McKinsey & Company. (2020). Supply chain risk management is back. https://www.mckinsey.com/capabilities/operations/our-insights/supply-chain-risk-management-is-back

Melo, M. T., Nickel, S., & Saldanha-da-Gama, F. (2009). Facility location and supply chain management – A review. *European Journal of Operational Research, 196*(2), 401–412. https://doi.org/10.1016/j.ejor.2008.05.007

Rinaldi, M., Murino, T., Gebennini, E., Morea, D., & Bottani, E. (2022). A literature review on quantitative models for supply chain risk management: Can they be applied to pandemic disruptions? *Computers & Industrial Engineering, 170,* 108329. https://doi.org/10.1016/j.cie.2022.108329

Schleper, M. C., Duensing, S., & Busse, C. (2024). Shifting the context: reviews and research agendas for traditional, reputational and societal supply chain risk. Supply Chain Management: *An International Journal, 29*(7). https://doi.org/10.1108/scm-04-2024-0280

Schmitt, A. J., & Singh, M. (2009). Quantifying supply chain disruption risk using Monte Carlo and discrete-event simulation. *Proceedings of the 2009 Winter Simulation Conference (WSC),* 1237–1248. https://doi.org/10.1109/wsc.2009.5429561

Sheffi, Y., & Rice, J. B. (2005, October 15). *A supply chain view of the resilient enterprise.* MIT Sloan Management Review. https://sloanreview.mit.edu/article/a-supply-chain-view-of-the-resilient-enterprise/

Singhal, P., Agarwal, G., & Mittal, M. L. (2011). Supply chain risk management: Review, classification and future research directions. *International Journal of Business Science and Applied Management, 6*(3), 15–52. https://doi.org/10.69864/ijbsam.6-3.75

Snyder, L. V., Atan, Z., Peng, P., Rong, Y., Schmitt, A. J., & Sinsoysal, B. (2016). OR/MS models for supply chain disruptions: a review. *IIE Transactions, 48*(2), 89–109. https://doi.org/10.1080/0740817x.2015.1067735

Tang, C. S. (2006). Perspectives in supply chain risk management. *International Journal of Production Economics, 103*(2), 451–488. https://doi.org/10.1016/j.ijpe.2005.12.006

Wang, W., Hocke, K., & Mätzler, C. (2021). Physical retrieval of rain rate from ground-based microwave radiometry. *Atmosphere, 12*(3), 364. https://doi.org/10.3390/atmos12030364

Wieland, A. (2021). Dancing the Supply Chain: toward Transformative Supply Chain Management. *Journal of Supply Chain Management, 57*(1). https://doi.org/10.1111/jscm.12248