CIRRELT

CIRRELT-2026-23

Tactical Planning of an Integrated Multi-stakeholder Freight Transport System under Uncertainty in Demand and Supply

Nadia Tahmasbi
Teodor Gabriel Crainic
Walter Rei
Janosch Ortmann

Novembre 2026

Bureau de Montréal
Université de Montréal
C.P. 6128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

Tactical Planning of an Integrated Multi-stakeholder Freight Transport System under Uncertainty in Demand and Supply

Nadia Tahmasbi[1,2,3], Teodor Gabriel Crainic[1,2*], Walter Rei[1,2], Janosch Ortmann[1,3]

1. School of Management, Université du Québec à Montréal, Montréal, Canada.
2. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), Montréal, Canada.
3. Group for Research in Decision Analysis (GERAD), Montréal, Canada.

Abstract. We consider the tactical planning of an integrated multi-stakeholder freight transportation system in which an Intelligent Decision Support Platform selects shipment requests from shippers and service offers from carriers and plans shipment itineraries. At the tactical level, the estimated volumes of the regular requests that will be received are subject to fluctuations, and because the platform relies on capacity it does not own, the estimated regular capacities of the services are uncertain as well. To account for both sources of uncertainty, we model the problem as a two-stage stochastic scheduled service network design problem: requests, services, and itineraries are selected in the first stage, and the execution of the plan is adjusted to the realized volumes and capacities in the second stage. We prove that the recourse problem is equivalent to a multidimensional knapsack problem, and adapt a decision-based scenario clustering method to bound the optimal value of instances too large to be solved directly. The computational study first evaluates the tractability of the formulation and the quality of the bounds. It then quantifies the value of the stochastic solution and the relative impact of demand and supply uncertainty. Finally, it compares the stochastic and deterministic plans and examines the setting of the non-fulfillment penalty. The results show that accounting for uncertainty yields more profitable plans, and an increase in capacity variability raises this gain about 15 times as much as the same increase in demand variability. The stochastic plans obtain their advantage by accepting fewer non-contract and irregular shipments and relying more on contract carriers, while keeping the non-fulfillment penalty above a threshold maintains service levels at a small cost in profit.

Keywords: Transportation, Stochastic programming, Scheduled service network design, Integrated consolidation-based freight transportation, Supply and demand uncertainty

Acknowledgments: While working on the project, the second author held the UQAM Chair on Intelligent Logistics and Transportation Systems Planning and was Adjunct Professor, Department of Computer Science and Operations Research, Université de Montréal, while the third held the Canada Research Chair in Stochastic Optimization of Transport and Logistics Systems, UQAM. The authors gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Collaborative Research and Development, Discovery, Acceleration, and Engage grant programs. We also gratefully acknowledge the support of the Fonds de recherche du Québec through their CIRRELT infrastructure grants. Thanks are also due to Calcul Québec and Compute Canada for providing the authors access to their high-performance computing infrastructure.

Results and views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect those of CIRRELT.

Les résultats et opinions contenus dans cette publication ne reflètent pas nécessairement la position du CIRRELT et n'engagent pas sa responsabilité.

---

* Corresponding author: teodorgabriel.crainic@cirrelt.net

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2026

© Tahmasbi, Crainic, Rei, Ortmann and CIRRELT, 2026

1 Introduction

The transportation and logistics (TL) sector is a cornerstone of the global economy, facilitating the seamless movement of goods and services across regions and countries. Not only does it support national and international trade, but it also contributes significantly to economic growth and job creation. According to the U.S. Department of Transportation, in 2023, the U.S. transportation system moved an average of about 55.5 million tons of freight per day, valued at more than $51.2 billion, and tonnage is estimated to increase by approximately 1.2 percent per year between 2024 and 2050 (Bureau of Transportation Statistics). As global markets expand and consumer demands evolve, there is increasing pressure on the TL sector to achieve greater efficiency and profitability while also reducing the environmental impacts that stem directly from the sector's activities. Technologies, particularly advanced analytical methods such as operations research, are playing a key role in this transformation, as illustrated by the development of intelligent freight transportation systems and smart logistics approaches that support higher levels of efficiency in the planning and execution of TL operations (Crainic et al., 2009; Feng and Ye, 2021).

In this study, we consider an integrated multi-stakeholder freight transportation system in which *many* shippers, such as retailers, distributors, and manufacturers, submit *shipment requests* to be moved from their origins to their destinations in a cost- and time-efficient manner. On the other side, *many* carriers, such as freight carriers, terminal managers, and logistics service providers, submit *capacity offers* for transportation services and warehousing space while seeking profitable shipments. In the middle, an *Intelligent Decision Support Platform (IDSP)* performs the planning and optimization of operations to profitably and simultaneously satisfy the needs of both parties. From a system-design perspective, this platform exhibits a Many-to-One-to-Many structure; thus, we refer to this system as an M1M system. Such systems include digital platforms operated by logistics service providers, as well as multi-stakeholder systems organized according to City logistics, Physical Internet, and synchromodal principles.

The IDSP plays a crucial role in ensuring the success of such a platform, as it supports decision-making across three interconnected levels: strategic, tactical, and operational. At the strategic level, it oversees long-term decisions, such as selecting strategic partners among shippers and carriers or investing in infrastructure and technology. At the tactical level, the IDSP focuses on medium-term decisions, including the selection of both shipment requests and capacity offers, which include transportation services and warehousing space, as well as the assignment of shipments to selected capacity offers and the specification of shipment itineraries that may combine transportation services with temporary warehousing to enable requests to be met. These tactical decisions are essential for bridging strategic objectives and day-to-day operations. Finally, at the operational level, the IDSP revises its tactical decisions, at least partially, to adapt to real-time conditions. *In this study, we focus exclusively on the tactical planning of the M1M system.*

The tactical plan is designed to be repeatedly executed over a medium-term planning horizon, referred to as *a season* (Crainic, 2024). In the context of the M1M system, this tactical-planning setting differs from more classical shipper- or carrier-oriented settings, as it must simultaneously account for a significant volume of *regular demand*, i.e., recurring and somewhat predictable shipment requests, and *regular supply*, i.e., capacity offers that can be committed over the considered horizon and are also somewhat predictable in nature. The plan is developed for a specific time duration, known as the *schedule length*, e.g., one week, and applied repeatedly throughout the season. The primary goal of tactical planning in the M1M system is to select the shipment requests to serve and to design the service network and shipment itineraries that maximize platform profitability while satisfying shipper and carrier requirements.

A distinguishing feature of the M1M system is the heterogeneity of its stakeholders and the different types of relationships they maintain with the platform. On the demand side, shipment requests are categorized according to the shippers' relationship with the platform: regular demand from contract-based shippers, whose requests are associated with commitments negotiated in advance; regular demand from non-contract shippers, whose recurring requests may be accepted or rejected by the platform during tactical planning; and irregular demand, which is estimated from historical data and treated as supplementary to the platform's core operations. On the supply side, capacity offers are similarly categorized according to the carriers' relationship with the platform: offers from contract-based carriers operating under long-term agreements; offers from non-contract carriers, which the platform evaluates during tactical planning; and ad hoc capacity obtained from spot markets to address urgent or unforeseen transportation needs.

Another important feature of the considered system is the informational uncertainty under which it is planned and operated. Specifically, on the demand side, the volumes associated with shipment requests may randomly fluctuate throughout the tactical horizon, regardless of the specific relationship that binds the corresponding shippers to the platform. Moreover, on the supply side, because the platform does not own transportation resources and depends on carriers to honor their capacity commitments, the availability of capacity offers may also be uncertain. Even contract-based carriers may fall short of agreed terms due to equipment failures, driver illness, or strikes, while non-contract carriers may modify or withdraw their services entirely. When tactical planning is performed by the IDSP for the M1M system, it is therefore essential to explicitly consider both demand- and supply-side uncertainty. These two sources of uncertainty jointly affect the reliability of the tactical plan by increasing the risk that selected shipment requests, retained capacity offers, and planned shipment itineraries may no longer be feasible or profitable when the plan is executed.

The emergence of such integrated freight platforms has attracted growing attention in both industry and the logistics literature (Kam et al., 2026). The tactical planning of integrated, consolidation-based systems raises challenging modeling and algorithmic

issues, with the joint treatment of demand, supply, and the uncertainty affecting both emerging as a central direction for further study (Crainic and Rei, 2025).

Recently, Taherkhani et al. (2022) proposed a scheduled service network design formulation for a tactical planning problem in an M1M system, incorporating revenue management concepts such as shipper categories, demand classification, penalty costs, and service bundles. However, their approach relies on point forecasts for both freight volumes and service capacity offers. A subsequent study by Taherkhani et al. (2025) extended this framework by introducing demand-side uncertainty. While this extension accounts for one important source of informational uncertainty in M1M systems, supply-side uncertainty remains unaddressed in this tactical-planning context. This omission can be significant in platform-based systems, as it may expose the IDSP to risks that are not accounted for when tactical plans are designed.

To address this gap, we develop a Scheduled Service Network Design (SSND) formulation that explicitly accounts for variability in both demand volumes and service capacities. Specifically, we make the following contributions:

(i) We formulate the SSND under Uncertainty in Demand and Supply (SSND-UDS), a two-stage stochastic program in which first-stage decisions (service selection, shipment acceptance, and tactical itinerary construction) are made before uncertainty is resolved, while second-stage recourse decisions adapt these itineraries to each realized scenario of shipment volumes and service capacities.

(ii) We show that the second-stage recourse problem admits an equivalent multidimensional knapsack problem (MKP) reformulation, which enables an efficient evaluation of fixed first-stage decisions.

(iii) For instances where solving the full stochastic program becomes computationally prohibitive, particularly when the sample-based approximation used for computation involves many demand-and-supply scenarios and many shipments, we adapt the decision-based scenario clustering approach of Hewitt et al. (2022) to provide high-quality lower and upper bounds on the optimal value, thereby enabling tractable planning when direct solution is infeasible within practical time limits.

(iv) Through computational experiments on a comprehensive set of benchmark instances, we show that explicitly accounting for supply-side uncertainty is necessary for the tactical planning of M1M systems: supply-side uncertainty contributes approximately 15 times more to the Value of the Stochastic Solution (VSS) than demand-side uncertainty, implying that tactical plans optimized against demand uncertainty alone capture only a small fraction of the total value of stochastic planning. We further show the operational benefits of stochastic planning in terms of tactical-plan robustness and recourse-cost reduction across shipper categories and uncertainty regimes.

The paper is organized as follows. Section 2 reviews the relevant literature on service network design, revenue management in freight transportation, stochastic network design, and scenario clustering. Section 3 presents the problem setting, including the physical network, the stakeholder structure, and the tactical planning process under uncertainty. The two-stage stochastic formulation is developed in Section 4. Section 5 describes the solution approach: the direct solution of the extensive form with Gurobi, the MKP reformulation of the recourse problem, and the scenario clustering bounds for larger instances. Computational experiments and managerial insights are reported in Section 6, and Section 7 concludes with a summary of findings, managerial insights, and directions for future research.

2 Literature review

The SSND-UDS problem integrates elements from several streams of the operations research and logistics literature. This section reviews the most relevant contributions along four themes: service network design for consolidation-based carriers, revenue management in freight transportation, stochastic (service) network design, and employing scenario clustering solution methods for solving network design problems.

Service network design (SND) is the foundational framework underlying our model. In SND, the goal is to jointly determine the set of services to operate and the routing of commodities through those services to minimize cost or maximize profit. The study by Crainic (2000) establishes the conceptual and algorithmic foundations of SND for consolidation-based freight carriers, covering formulations for less-than-truckload, railway, intermodal, and multi-commodity settings. Crainic and Rei (2025), provides a comprehensive survey of five decades of operations research contributions to consolidation-based freight tactical planning, tracking the evolution from static SND formulations to time-sensitive Scheduled Service Network Design (SSND) models and, ultimately, to stochastic extensions that explicitly account for uncertainty in demand and supply.

Tactical planning for consolidation-based freight carriers has been widely addressed through the SND problems across a variety of transportation modes and contexts. Please refer to Crainic (1999) for general long-haul freight transportation, Christiansen et al. (2007) and Christiansen et al. (2020) for maritime transportation, Chouman and Crainic (2021), and Cordeau et al. (1998) for rail systems, Crainic and Kim (2007) and Steadie-Seifi et al. (2014) for intermodal networks, Bakir et al. (2021) for motor carriers.

Despite the extensive SND literature, profit-maximizing formulations for freight transportation planning remain relatively rare. The majority of the studies that incorporates revenue management concepts has focused on the operational level, where decisions are made closer to execution. We refer the reader to Crevier et al. (2012), Wang et al. (2014),

Bilegan et al. (2015), Wang et al. (2016), van Riessen et al. (2017), and Van Riessen et al. (2021) for studies that incorporate revenue management considerations into carrier service planning with the aim of maximizing profitability under operational conditions. At the tactical level, Bilegan et al. (2022) examine SSND for intermodal barge transportation under profit-maximizing objectives, incorporating differentiated customer categories and tariff structures to maximize net carrier revenue. Building on this line of work, Taherkhani et al. (2022) address the tactical planning problem within the M1M system by integrating shipper categorization, demand classification, time-window penalty structures for late or missed deliveries, and service bundle offerings. Their study develops and compares alternative deterministic formulations of the problem, evaluating computational performance using the CPLEX solver.

Most studies in the SSND literature adopt deterministic settings; incorporating uncertainty into service network design problems is a more recent and steadily growing line of research (Klibi et al., 2010), (Hewitt et al., 2021). Among the various sources of uncertainty considered in the literature, demand uncertainty has received the most attention. Lium et al. (2009) use a two-stage stochastic programming formulation to show that explicitly modeling demand uncertainty produces qualitatively different network structures: consolidation and hub-and-spoke patterns emerge organically as hedging mechanisms, and demand correlation structures significantly influence the resulting plan. Their computational results show that schedules derived from stochastic models achieve substantially lower expected costs compared to those obtained from deterministic approaches. Hewitt et al. (2019) address SSND with resource acquisition and management under demand uncertainty, proposing a two-stage scenario-based stochastic model solved via a column-generation-based matheuristic.

Beyond demand uncertainty, several studies incorporate multiple sources of stochasticity simultaneously. Zhao et al. (2018) formulate a two-stage chance-constrained program for intermodal sea-rail service network design under stochastic travel times, transfer times, and container demand, with on-time delivery and capacity chance constraints handled by a hybrid heuristic. Guo et al. (2021) study a dynamic and stochastic shipment matching problem on a logistics platform in synchromodal transportation, where both shipment request attributes (origin, destination, volume, time window, and fare class) and service travel times are uncertain. Their rolling horizon framework with a hybrid stochastic approach shows that explicitly modeling multiple uncertainty sources yields significant gains in profitability and delivery reliability compared to approaches that ignore or partially capture uncertainty.

Most closely related to our work, Taherkhani et al. (2025) extend the M1M tactical planning framework of Taherkhani et al. (2022) to account for demand-side uncertainty. They develop a two-stage stochastic SSND formulation in which shipment volume uncertainty is modeled through scenarios, and propose a Benders decomposition-based algorithm with several acceleration techniques to solve the resulting model. Their

computational analysis quantifies the value of stochastic planning over deterministic approaches and provides insights into how uncertainty affects service network structure and platform profitability. However, their model assumes that service capacities offered by carriers are deterministic, an assumption that overlooks a practically significant source of risk. Platform operators depend entirely on carriers to deliver promised capacities; even contract-based carriers may fall short due to equipment failures, strikes, or operational disruptions, while non-contract carriers face additional incentive misalignment. To the best of our knowledge, no existing study addresses tactical planning in the M1M system under simultaneous uncertainty in both demand volumes and service capacities, a limitation that motivates the present work, which extends the M1M stochastic planning framework to explicitly account for uncertainty on both the demand and supply sides.

A fundamental challenge in scenario-based stochastic programming is that the large number of scenarios required to faithfully represent uncertainty renders the resulting formulations computationally intractable, particularly for stochastic integer programs. This motivates the study of scenario clustering methods, which approximate the original scenario distribution with a smaller representative subset while limiting the error introduced.

The dominant class of approaches is *distribution-driven*: methods such as k-means, k-medoids, and Monte Carlo sub-sampling select representative scenarios by minimizing a metric distance in the space of random outcomes, without reference to the optimization model. While computationally accessible, these methods may cluster scenarios that, despite being close in distribution space, induce very different first-stage decisions.

To address this, Keutchayan et al. (2021) propose the Cost-Space Scenario Clustering (CSSC) algorithm, a *problem-driven* approach that clusters scenarios according to their values in the objective function space rather than in the space of random outcomes. The method minimizes a discrepancy measure that bounds the implementation error, i.e., the loss incurred by applying a solution derived from a reduced scenario set to the full problem. Applied to stochastic network design and facility location problems, CSSC substantially outperforms distribution-driven alternatives, demonstrating that representative scenarios need not reproduce the moments of the original distribution to yield high-quality solutions.

Building on this problem-driven perspective, Hewitt et al. (2022) introduce a *decision-based scenario clustering* framework that measures dissimilarity between scenarios through the opportunity cost incurred when a first-stage plan optimized for one scenario is applied when another scenario is realized. This decision-driven distance is then used to partition scenarios via graph-based clustering, grouping scenarios that induce mutually acceptable decisions. Applied to stochastic network design and fleet planning problems, the approach yields high-quality lower and upper bounds at a fraction of the cost of direct solution.

In this paper, we adopt the decision-based distance of Hewitt et al. (2022) combined with the Partitioning Around Medoids algorithm (Keutchayan et al., 2021) to cluster scenarios for the SSND-UDS problem. Because our uncertainty spans both demand volumes and service capacities simultaneously, the decision-based distance naturally captures the joint impact of both sources on platform profitability, without requiring an arbitrary weighting between the two dimensions as distribution-driven measures would.

3 Problem Setting

In this section, we describe the general freight transportation system considered and present the tactical planning process it entails, thereby directly motivating the SSND-UDS proposed to address it. In Section 3.1, we define the role played by the IDSP within the M1M system and introduce the physical transportation network over which its activities take place. In Section 3.2, we characterize the heterogeneous stakeholders on both the demand and supply sides. Lastly, in Section 3.3, we formalize the tactical planning process, while also defining the sources of uncertainty that are explicitly considered and that directly motivate the stochastic formulation.

3.1 The IDSP and the Physical Transportation Network

In the middle of the M1M system, an Intelligent Decision Support Platform automates the planning and optimization of operations to profitably and simultaneously satisfy the needs of both shippers and carriers. It coordinates and plans both sides of the system: it determines which shipper requests should be served and which carrier service offers should be retained to provide the transportation capacity required to satisfy this demand. These two sets of decisions cannot be taken separately, since a request is worth serving only if capacity can be secured to move it within its time limits, and a service offer is worth retaining only if the shipments assigned to it generate enough revenue to cover its cost.

Although the platform earns its revenues from the shipper side rather than from the carriers, its planning decisions are largely structured around the carrier side of the system. The IDSP seeks to select and serve profitable shipment requests, but it does not own transportation assets; service capacity is entirely sourced from carriers through contractual or spot-market agreements that must be secured before operations begin. The revenue that a request can generate is therefore known with reasonable confidence when the plan is built, whereas the means of realizing it remain in the hands of independent carriers.

The IDSP manages decision-making across three interconnected levels. At the strategic level, it oversees long-term decisions such as selecting partners and investing in infrastructure. At the tactical level, it selects shipment requests and capacity offers, assigns shipments to services, and constructs shipment itineraries. At the operational level, it revises tactical decisions to adapt to real-time conditions. This study focuses exclusively on the tactical planning level.

The role played by the IDSP differs from that of a traditional consolidation-based carrier, such as a less-than-truckload motor carrier (Bakir et al., 2021) or an intermodal barge operator (Bilegan et al., 2022), as well as from that of an online marketplace matching shippers with transportation providers, such as the load boards operated by DAT Freight & Analytics and Truckstop. Unlike a carrier, the IDSP does not operate the services over which it consolidates freight: a carrier plans its network over capacity it controls, whereas the IDSP selects services from offers submitted by independent carriers, and the capacity of each selected service remains a commitment made by a third party. Unlike a marketplace, which enables shippers and carriers to connect and negotiate directly and leaves the routing of each shipment to the parties involved, the IDSP constructs the itinerary that each accepted shipment must follow through the network and answers to the shipper for its execution. The IDSP therefore faces the planning problem of a consolidation-based carrier while depending, as a marketplace does, on the capacity it neither owns nor controls.

Transport activities in the M1M system take place on a network of consolidation and transfer terminals. *Consolidation terminals* are located near geographic zones and handle first-mile and last-mile services, sorting, and cross-docking. *Transfer terminals* offer only sorting and cross-docking functions. Each terminal associated with a zone serves all origins and destinations within that zone, so pickup and delivery activities within zones are not explicitly modeled. On the demand side, shippers submit shipment requests to be transported freight between geographic zones; since each zone is served by a dedicated terminal, shipment origins and destinations are represented at the terminal level throughout the model. On the supply side, carriers offer capacity for transport services between terminals and storage spaces at terminals, to which we refer simply as *services*. Services connect terminals directly, each operating between a specific origin-destination pair without intermediate stops.

Figure 1 illustrates a representative instance of this network. Terminals A, C, and D are consolidation terminals, each associated with a geographic zone, while terminals B and E, shown in grey, are transfer terminals, which belong to no zone and allow freight originating at several consolidation terminals to be regrouped onto long-haul services and redistributed toward its destinations. Each line represents a service offered by a carrier between the two terminals it connects. Services may be operated by different modes (road, rail, and barge in the illustration), which differ in cost, capacity, and travel time; the model, however, treats them uniformly, as a service is characterized only by its

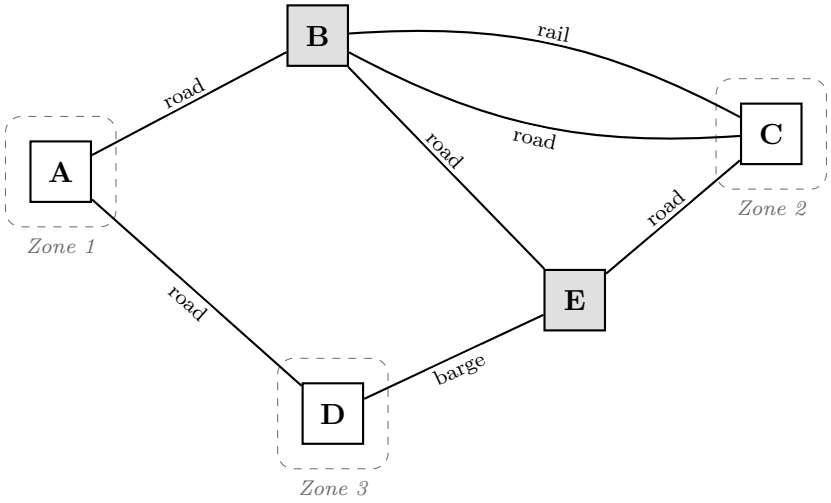


Figure 1: A representative instance of the physical transportation network. Consolidation terminals are shown in white and transfer terminals in grey.

origin and destination terminals, its departure and arrival periods, its capacity, and its cost structure. Several services may be offered on the same terminal pair, by different carriers, at different times, or using different modes, as between terminals B and C in the figure, and each is selected independently by the platform.

3.2 Stakeholder Structure

The IDSP interacts with two distinct groups of stakeholders: shippers, who generate freight demand, and carriers, who provide transportation capacity. Both groups are internally heterogeneous: stakeholders within each group differ in their contractual relationship with the platform, their fulfillment obligations, and the uncertainty characteristics of their demand or capacity. Table 1 summarizes the three shipper categories and three carrier categories along their key planning dimensions.

Each *shipment request* is characterized by its origin terminal, destination terminal, release time, delivery due time, and unit revenue. Pickups before the release time are not permitted, late pickups and late deliveries are allowed but incur penalties, while early deliveries are permitted without penalty. Individual shipments may not be split: once a request is accepted, the corresponding freight is assigned to a single tactical itinerary. Since the platform acts as an intermediary between shippers and carriers and does not directly control the transportation resources, the tactical plans it produces must remain simple to communicate, coordinate, and execute across independent stakeholders. A *tactical itinerary* is the plan established during tactical planning for transporting a shipment from its origin to its destination, specifying the services it uses, the terminals

Table 1: Stakeholder Categories and Planning Characteristics

	Contract-based	**Non-contract**	**Irregular / Ad hoc**
Shippers			
Request type	Regular demand; capacity secured via strategic partnership	Regular demand; submitted well in advance	Irregular demand; submitted on short notice; forecasted from historical data
Fulfillment obligation	Must be fulfilled under all circumstances	Must be fulfilled once accepted; non-fulfillment penalty applies	Supplementary; once accepted, lower non-fulfillment penalty applies
Volume uncertainty	Realized volume may differ from announced; no penalty imposed	Demand increases accepted based on profitability; decreases incur no penalty	Same as non-contract
Carriers			
Service type	Long-term agreements; most reliable services	Service offers submitted in advance; platform decides to accept or reject	Spot market; procured per shipment on short notice; higher unit cost
Cost structure	Fixed cost already incurred; unit cost only	Fixed cost + unit cost when selected	Unit cost only; no fixed cost
Capacity uncertainty	Capacity may fall short due to equipment failure, strikes, or disruptions	Higher uncertainty; carriers may modify or withdraw services	Not applicable; procured after uncertainty is realized

at which it is handled, and the periods in which it is picked up, transferred, and delivered.

The cost structure for carrier services is differentiated by category and reflects the nature of each carrier relationship. For contract-based carriers, fixed costs have already been incurred under long-term agreements, so only the unit cost for utilizing these services applies. Non-contract services incur both a fixed reservation cost and a variable unit cost when selected by the platform. Ad hoc services, procured from spot markets on a per-shipment basis, incur only a unit cost with no associated fixed component.

3.3 Tactical Planning Under Uncertainty

The tactical plan is designed to be repeatedly executed over a medium-term planning horizon, referred to as a season. The plan is developed for a specific time duration, known as the schedule length (e.g., one week), and applied repeatedly throughout the season. The primary goal of tactical planning in the M1M system is to design service networks and shipment routes that maximize platform profitability while meeting shipper and carrier requirements.

Since tactical planning takes place months before actual operations, both shipment volumes and service capacities are subject to uncertainty at the time plans are formulated. This variability, categorized as business-as-usual uncertainty (Klibi et al., 2010), is observed throughout the planning horizon and must be managed when the plan is implemented. The uncertainty characteristics by stakeholder category are summarized in Table 1.

As noted in Section 3.1, the IDSP is entirely dependent on the services provided and managed by various carriers to fulfill shipment requests. This dependence introduces an additional layer of operational uncertainty, as the IDSP depends on carriers to adhere to their announced schedules and to provide the contracted capacity. Deviations from these commitments can directly affect the IDSP's ability to meet customer demand and maintain service reliability. Even for contract-based services—the most reliable providers—actual capacities may fall short of agreed terms due to factors such as equipment failures, driver illness, traffic congestion, or strikes. The situation is even more uncertain for non-contract carriers, who lack long-term commitments to the platform and prioritize their own profitability. This increases uncertainty in service capacities, stemming from both unpredictable events and carriers' decisions to modify or withdraw services. Explicitly accounting for supply-side uncertainty in the tactical plan is therefore an operational necessity.

Beyond selecting requests and services, the IDSP coordinates the activities of the stakeholders it brings together. The timing and operational information contained in the itineraries is communicated to shippers and carriers in advance of the season, so that each

may organize its own activities accordingly: shippers arrange the release of their freight at the announced periods, and carriers plan their operations around the loads they are expected to carry on each departure. The tactical plan is thus not merely an internal planning tool for the platform, but a coordination and communication mechanism across the system.

To precisely define the decision-making process, we distinguish between decisions made before uncertainty is resolved, and the decisions made afterward (i.e., the recourse). First-stage decisions encompass service selection, shipment acceptance, and tactical itinerary construction. These decisions are made before shipment volumes and service capacities are realized. Second-stage recourse decisions adapt the execution of pre-committed itineraries to each scenario.

Recourse actions in stochastic programming range from simple recourse (observing outcomes and incurring associated costs) to full network re-optimization, where all first-stage decisions are revisited. The coordination role described above determines where our formulation sits on this spectrum. The services and tactical itineraries established in the first stage remain fixed across scenarios, while only the decisions on whether and how to satisfy each shipment (via the tactical itinerary, an ad hoc service, or non-fulfillment) are scenario-dependent. Both stages are solved simultaneously. Importantly, second-stage decisions are incorporated not to plan specific operational itineraries, but to ensure that the first-stage tactical plan performs effectively across all demand and supply scenarios.

In the second-stage, several cases may arise due to the realization of shipment volumes and service capacities. If the total volume of shipments assigned to a service exceeds realized capacity, we refer to that service as an overflow service. For contract-based shipments, the platform is obligated to fulfill the request regardless of overflow, and outsources delivery to an ad hoc carrier when the tactical itinerary cannot be executed. For non-contract and irregular shipments, the platform may either outsource delivery to an ad hoc carrier or forgo fulfillment and incur the corresponding non-fulfillment penalty. This decision is guided by profitability, given the higher unit cost of ad hoc services relative to planned itineraries.

4 Optimization Model

Building on the problem setting described in Section 3, this section develops the mathematical formulation of the SSND-UDS problem. We begin by formalizing the time-space network representation used to model the system (Section 4.1). Such a representation is required because tactical plans prescribe how scheduled transportation services and temporary warehousing activities are deployed over the schedule length, while commodity itineraries must specify feasible paths through both space and time to ensure that ship-

ment requests can be delivered within the required time windows. We then introduce the notation (Section 4.2) and present the two-stage stochastic program (Section 4.3), which is formulated over the joint distribution of shipment volumes and service capacities. The sample-based approximation through which the model is solved is introduced in Section 5, and the distributions from which the sample is drawn are described in Section 6.

4.1 Time-Space Network Representation

Transport decisions in the M1M system are modeled on a time-space network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ constructed by expanding the physical network $\mathcal{G}^{PH} = (\mathcal{N}^{PH}, \mathcal{A}^{PH})$ over the schedule length. The schedule length is divided into $|\mathscr{T}|$ periods of equal length, set exogenously prior to planning, and $\mathscr{T} = \{1, 2, \ldots, |\mathscr{T}|\}$ denotes the resulting set of time periods. For each terminal $\eta \in \mathcal{N}^{PH}$, we let $\mathscr{T}(\eta) \subseteq \mathscr{T}$ denote the periods at which an event can occur at η. Time periods are retained for a terminal only when an event can occur there; for instance, no time-space node is created for a shipment's origin terminal prior to its release time. The node set of $\mathcal{G}$ is thus $\mathcal{N} = \{(\eta, t) \mid \eta \in \mathcal{N}^{PH},\ t \in \mathscr{T}(\eta)\}$.

The arc set $\mathcal{A} = \mathcal{A}^M \bigcup \mathcal{A}^H$ comprises two types. *Moving arcs*, $\mathcal{A}^M$, represent direct transportation between geographically distinct terminals: each arc $a = ((o_\sigma, \alpha_\sigma), (d_\sigma, \beta_\sigma))$ corresponds to a service $\sigma \in \Sigma$ departing terminal o_σ at time α_σ and arriving at terminal d_σ at time β_σ. Following Taherkhani et al. (2025), we assume that carriers offer single-leg services, that is, services that move freight directly between two terminals, with no service operated over a path involving three or more terminals. This assumption entails no significant loss of generality, since a multi-leg service can be represented as a sequence of single-leg services, one per leg, provided that the corresponding arcs are selected jointly and that the fixed cost of the service is charged only once.

Holding arcs, $\mathcal{A}^H$, represent the temporary storage of freight at a terminal for a single period: $a = ((\eta, t), (\eta, t+1))$ for $\eta \in \mathcal{N}^{PH}$ and $t, t+1 \in \mathscr{T}(\eta)$. Since there is a one-to-one correspondence between arcs and offered services, each arc a inherits the attributes of its corresponding service: origin o_a, destination d_a, fixed cost f_a, and shipment-specific unit cost c_a^k. For each time-space node (η, t), we let $\mathcal{A}^+(\eta, t)$ and $\mathcal{A}^-(\eta, t)$ denote, respectively, the sets of arcs departing from and arriving at (η, t). Figure 2 illustrates a simple time-space network and its corresponding physical network, where each node is duplicated for all five time periods.

4.2 Sets, Parameters, and Decision Variables

Tables 2 and 3 summarize the sets and parameters, and the decision variables of the model, respectively. The transportation services and shipment requests introduced in

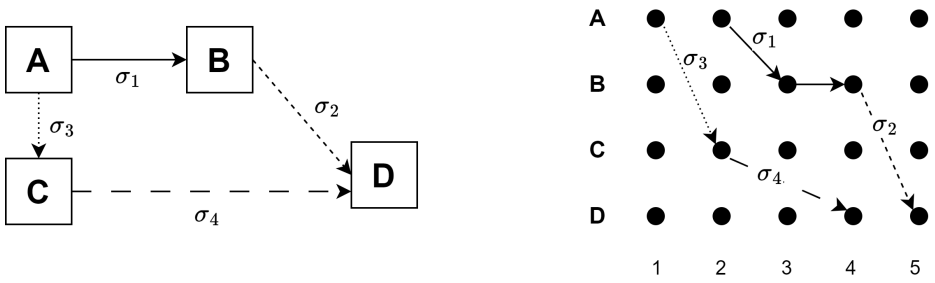


Figure 2: Physical network and its corresponding time-space network

Section 3 are represented using the following partitions: $\Sigma = \Sigma^C \cup \Sigma^N \cup \Sigma^A$ and $\mathcal{K} = \mathcal{K}^C \cup \mathcal{K}^N \cup \mathcal{K}^I$. For transportation services, the superscripts C, N, and A refer to contract-based, non-contract, and ad hoc services, respectively. For shipment requests, the superscripts C, N, and I refer to contract-based, non-contract, and irregular requests, respectively.

Cost structures also differ across service categories. For contract-based services, fixed costs have already been incurred under long-term agreements; therefore, we assume $f_\sigma = 0$, for all $\sigma \in \Sigma^C$. Ad hoc services incur only a unit cost, with no fixed component, so that $f_\sigma = 0$ and $c_\sigma^k > 0$, for all $\sigma \in \Sigma^A(k)$. Finally, non-contract services incur both a fixed and a variable cost.

4.3 Two-Stage Stochastic Formulation

The SSND-UDS is formulated as a two-stage stochastic mixed-integer program. As established in Section 3, first-stage decisions are made before uncertainty is realized and encompass service selection, shipment acceptance, and tactical itinerary construction; second-stage recourse decisions adapt these itineraries to the realized shipment volumes and service capacities.

Let ξ denote the random vector collecting the uncertain volumes $w_k(\xi)$ of the shipment requests $k \in K$ and the uncertain capacities $u_a(\xi)$ of the arcs $a \in \mathcal{A}$, defined on the probability space $(\Xi, \mathcal{F}, \mathbb{P})$, where Ξ denotes the support of ξ. Ad hoc services are procured from the spot market after the uncertainty is realized, at known unit costs, and are therefore not subject to uncertainty; they remain available whatever the realization of ξ. First-stage decisions are taken before ξ is observed, while the second stage adapts the execution of the tactical plan to its realization.

First stage. The model selects non-contract services (y_σ), accepts shipment requests from non-contract and irregular shippers (z_k), and assigns all accepted shipments to moving and holding arcs ($x_{k,a}$ and $v_{k,a}$) , thereby establishing tactical itineraries. The

Table 2: Table of Sets and Parameters

Sets	**Description**
$\mathcal{K}$	Set of all shipment requests
$\mathcal{K}^C$	Set of regular shipment requests from contract-based shippers
$\mathcal{K}^N$	Set of regular shipment requests from non-contract shippers
$\mathcal{K}^I$	Set of irregular shipment requests from non-contract shippers
Σ	Set of all services
Σ^C	Set of services provided by contract-based carriers
Σ^N	Set of services provided by non-contract carriers
$\Sigma^A(k)$	Set of ad hoc services for shipment k obtained from spot market
$\mathcal{A}$	Set of arcs in the time-space network
$\mathcal{A}^M$	Set of arcs representing moving services
$\mathcal{A}^H$	Set of arcs representing holding services
$\mathcal{N}^{PH}$	Set of physical nodes in the network
$\mathcal{N}$	Set of time-space nodes
$\mathscr{T}(\eta)$	Set of time periods for terminal η
$\mathscr{T}$	Set of all time periods
Ξ	Support of the random vector ξ
Parameters	**Description**
o_k	Origin terminal of shipment k
d_k	Destination terminal of shipment k
α_k	Release time of shipment k at its origin terminal
β_k	Due time for delivery of shipment k at its destination terminal
ϕ_k	Penalty cost for failing to deliver shipment k
ρ_k	Revenue associated with satisfying shipment k
$\psi^P_{k,t}$	Penalty cost for late pickup of shipment k at time t
$\psi^D_{k,t}$	Penalty cost for late delivery of shipment k at time t
f_σ	Fixed cost of selecting service σ
c^k_σ	Unit cost of using service σ for shipment k
t^η_l	Start time of time period l for terminal η
t^η_{l+1}	End time of time period l for terminal η
o_σ	Origin terminal of service σ
d_σ	Destination terminal of service σ
α_σ	Departure time of service σ
β_σ	Arrival time of service σ
ξ	Random vector of shipment volumes and service capacities, $\xi \in \Xi$
$u_a(\xi)$	Capacity of arc a under realization ξ, corresponding to a service
$w_k(\xi)$	Volume of shipment k under realization ξ

Table 3: Decision variables, all binary

Variable	Description
First-stage variables	
$r^o_{k,t}$	Whether shipment request k is picked up from its origin o_k at time t.
$r^d_{k,t}$	Whether shipment request k is delivered to its destination d_k at time t.
y_σ	Whether service $\sigma \in \Sigma^N$ is selected.
z_k	Whether shipment request $k \in \mathcal{K}^N \cup \mathcal{K}^I$ is accepted.
$x_{k,a}$	Whether shipment request k is traveling on arc a.
$v_{k,a}$	Whether shipment request k is held on arc a.
Second-stage variables	
$l_k(\xi)$	Whether the tactical itinerary is used to satisfy shipment k under realization ξ.
$h_{k,\sigma}(\xi)$	Whether ad hoc service $\sigma \in \Sigma^A(k)$ is used to satisfy shipment k under realization ξ.
$e_k(\xi)$	Whether shipment $k \in \mathcal{K}^N \cup \mathcal{K}^I$ is satisfied under realization ξ.

binary variables $r^o_{k,t}$ and $r^d_{k,t}$ indicate whether shipment k is picked up at its origin or delivered to its destination at time period t, respectively.

Second stage. Given a realization of ξ, the second-stage problem determines, for each shipment, whether to execute the tactical itinerary ($l_k(\xi) = 1$), divert to an ad hoc service ($h_k(\xi) = 1$), or, for non-contract and irregular shipments, forgo fulfillment ($e_k(\xi) = 0$). Contract-based shipments must be fulfilled under all circumstances, via either the tactical itinerary or ad hoc services.

The formulation is as follows:

$$\max \quad \sum_{k\in\mathcal{K}^C} \rho_k - \sum_{\sigma\in\Sigma^N} f_\sigma y_\sigma + \mathbb{E}_\xi[Q(y,z,x,v,r;\xi)] \tag{1}$$

$$\text{s.t.} \quad \sum_{t\in\mathscr{T}(o_k)} r^o_{k,t} = 1, \qquad \forall k \in \mathcal{K}^C \tag{2}$$

$$\sum_{t\in\mathscr{T}(d_k)} r^d_{k,t} = 1, \qquad \forall k \in \mathcal{K}^C \tag{3}$$

$$\sum_{t\in\mathscr{T}(o_k)} r^o_{k,t} = z_k, \qquad \forall k \in \mathcal{K}^N \cup \mathcal{K}^I \tag{4}$$

$$\sum_{t\in\mathscr{T}(d_k)} r^d_{k,t} = z_k, \qquad \forall k \in \mathcal{K}^N \cup \mathcal{K}^I \tag{5}$$

$$r^o_{k,t} + v_{k,((o_k,t-1),(o_k,t))} = \sum_{a\in\mathcal{A}^+(o_k,t)} x_{k,a} + v_{k,((o_k,t),(o_k,t+1))},$$
$$\forall k \in \mathcal{K},\ \forall (o_k,t) \in \mathcal{N} \tag{6}$$

$$r^d_{k,t} + v_{k,((d_k,t),(d_k,t+1))} = v_{k,((d_k,t-1),(d_k,t))} + \sum_{a\in\mathcal{A}^-(d_k,t)} x_{k,a},$$
$$\forall k \in \mathcal{K},\ \forall (d_k,t) \in \mathcal{N} \tag{7}$$

$$\sum_{a\in\mathcal{A}^-(\eta,t)} x_{k,a} + v_{k,((\eta,t-1),(\eta,t))} = \sum_{a\in\mathcal{A}^+(\eta,t)} x_{k,a} + v_{k,((\eta,t),(\eta,t+1))},$$
$$\forall k \in \mathcal{K},\ \forall (\eta,t) \in \mathcal{N} \tag{8}$$

$$y_\sigma \in \{0,1\}, \qquad \forall \sigma \in \Sigma^N \tag{9}$$

$$z_k \in \{0,1\}, \qquad \forall k \in \mathcal{K}^N \cup \mathcal{K}^I \tag{10}$$

$$x_{k,a} \in \{0,1\}, \qquad \forall k \in \mathcal{K}, \forall a \in \mathcal{A}^M \tag{11}$$

$$v_{k,a} \in \{0,1\}, \qquad \forall k \in \mathcal{K}, \forall a \in \mathcal{A}^H \tag{12}$$

$$r^o_{k,t} \in \{0,1\}, \qquad \forall k \in \mathcal{K}, \forall (o_k,t) \in \mathcal{N} \tag{13}$$

$$r^d_{k,t} \in \{0,1\}, \qquad \forall k \in \mathcal{K}, \forall (d_k,t) \in \mathcal{N} \tag{14}$$

where

$$
\begin{aligned}
Q(x,y,z,r,v;\xi) = \max \quad & \sum_{k\in\mathcal{K}^N\cup\mathcal{K}^I} \rho_k e_k(\xi) \\
& - \sum_{k\in\mathcal{K}} \sum_{t\in\mathscr{T}(o_k)} \psi^P_{k,t} r^o_{k,t} l_k(\xi) - \sum_{k\in\mathcal{K}} \sum_{t\in\mathscr{T}(d_k)} \psi^D_{k,t} r^d_{k,t} l_k(\xi) \\
& - \sum_{k\in\mathcal{K}} \sum_{a\in\mathcal{A}^M} w_k(\xi) c^k_a x_{k,a} l_k(\xi) - \sum_{k\in\mathcal{K}} \sum_{a\in\mathcal{A}^H} w_k(\xi) c^k_a v_{k,a} l_k(\xi) \\
& - \sum_{k\in\mathcal{K}^N\cup\mathcal{K}^I} \phi_k(1-e_k(\xi)) z_k - \sum_{k\in\mathcal{K}} \sum_{\sigma\in\Sigma^A(k)} c^k_\sigma h_{k,\sigma}(\xi) \qquad (15)
\end{aligned}
$$

$$
\begin{aligned}
s.t. \quad & \\
& \sum_{\sigma\in\Sigma^A(k)} h_{k,\sigma}(\xi) + l_k(\xi) = 1, \qquad \forall k \in \mathcal{K}^C, \qquad (16) \\
& \sum_{\sigma\in\Sigma^A(k)} h_{k,\sigma}(\xi) + l_k(\xi) = e_k(\xi), \qquad \forall k \in \mathcal{K}^N \cup \mathcal{K}^I, \qquad (17) \\
& e_k(\xi) \leqslant z_k, \qquad \forall k \in \mathcal{K}^N \cup \mathcal{K}^I, \qquad (18) \\
& \sum_{k\in\mathcal{K}} w_k(\xi) x_{k,a} l_k(\xi) \leqslant u_a(\xi) y_a, \qquad \forall a \in \mathcal{A}^M, \qquad (19) \\
& \sum_{k\in\mathcal{K}} w_k(\xi) v_{k,a} l_k(\xi) \leqslant u_a(\xi) y_a, \qquad \forall a \in \mathcal{A}^H, \qquad (20) \\
& e_k(\xi) \in \{0,1\}, \qquad \forall k \in \mathcal{K}^N \cup \mathcal{K}^I, \qquad (21) \\
& l_k(\xi) \in \{0,1\}, \qquad \forall k \in \mathcal{K}, \qquad (22) \\
& h_{k,\sigma}(\xi) \in \{0,1\}, \qquad \forall k \in \mathcal{K}, \forall \sigma \in \Sigma^A(k). \qquad (23)
\end{aligned}
$$

The objective function (1) maximizes expected platform profit: the revenue from contract-based shipments is deterministic and enters the first stage, fixed service selection costs for non-contract carriers are subtracted, and the expected second-stage contribution is added, the expectation being taken over the joint distribution of shipment volumes and service capacities. Constraints (2) and (3) enforce mandatory pickup and delivery for contract-based shippers under all circumstances. In contrast, constraints (4) and (5) ensure that only accepted non-contract and irregular shipments are assigned pickup and delivery time periods. Constraints (6)-(8) enforce flow conservation for each shipment at its origin, destination, and intermediate terminals, respectively. Constraints (9)-(14) impose binary restrictions on all first-stage decision variables.

Equation (15) defines $Q(x,y,z,r,v;\xi)$ as the maximum profit under realization ξ given the first-stage decisions x,y,z,r,v. It includes the profit from satisfying non-contract shipments, minus penalties for late pickups and deliveries, and service costs when using tactical itineraries. It also accounts for penalties for failing to satisfy selected shipments and the costs of using ad hoc services. Constraint (16) ensures that shipments from contract-based shippers are either satisfied using tactical itineraries or assigned

to ad hoc services. For non-contract and irregular shippers, constraint (17) ensures that tactical itineraries or ad hoc services are used only if the shipment is satisfied. Constraint (18) ensures shipments are satisfied only if they are selected. Constraints (19) and (20) enforce capacity limits for moving and holding arcs, respectively. Finally, constraints (21)-(23) enforce binary requirements on second-stage decision variables.

The recourse function (15) and the capacity constraints (19) and (20) contain products of a first-stage binary variable with the second-stage variable $l_k(\xi)$. These products are linearized by introducing, for each shipment $k \in \mathcal{K}$ and each realization ξ, the auxiliary variables $\tilde{r}^o_{k,t}(\xi)$, $\tilde{r}^d_{k,t}(\xi)$, $\tilde{x}_{k,a}(\xi)$ and $\tilde{v}_{k,a}(\xi)$, which represent the products $r^o_{k,t} l_k(\xi)$, $r^d_{k,t} l_k(\xi)$, $x_{k,a} l_k(\xi)$ and $v_{k,a} l_k(\xi)$, respectively, and satisfy

$$\tilde{r}^o_{k,t}(\xi) \geqslant r^o_{k,t} + l_k(\xi) - 1, \qquad \forall k \in \mathcal{K},\ \forall t \in \mathscr{T}(o_k) \tag{24}$$
$$\tilde{r}^d_{k,t}(\xi) \geqslant r^d_{k,t} + l_k(\xi) - 1, \qquad \forall k \in \mathcal{K},\ \forall t \in \mathscr{T}(d_k) \tag{25}$$
$$\tilde{x}_{k,a}(\xi) \geqslant x_{k,a} + l_k(\xi) - 1, \qquad \forall k \in \mathcal{K},\ \forall a \in \mathcal{A}^M \tag{26}$$
$$\tilde{v}_{k,a}(\xi) \geqslant v_{k,a} + l_k(\xi) - 1, \qquad \forall k \in \mathcal{K},\ \forall a \in \mathcal{A}^H \tag{27}$$
$$\tilde{r}^o_{k,t}(\xi),\ \tilde{r}^d_{k,t}(\xi),\ \tilde{x}_{k,a}(\xi),\ \tilde{v}_{k,a}(\xi) \geqslant 0. \tag{28}$$

Each product in (15), (19) and (20) is replaced by the corresponding auxiliary variable, so that the penalty and transportation cost terms become $\psi^P_{k,t}\tilde{r}^o_{k,t}(\xi)$, $\psi^D_{k,t}\tilde{r}^d_{k,t}(\xi)$, $w_k(\xi)c^k_a\tilde{x}_{k,a}(\xi)$ and $w_k(\xi)c^k_a\tilde{v}_{k,a}(\xi)$, while the capacity constraints become $\sum_{k\in\mathcal{K}} w_k(\xi)\tilde{x}_{k,a}(\xi) \leqslant u_a(\xi)y_a$ for $a \in \mathcal{A}^M$ and $\sum_{k\in\mathcal{K}} w_k(\xi)\tilde{v}_{k,a}(\xi) \leqslant u_a(\xi)y_a$ for $a \in \mathcal{A}^H$.

Every auxiliary variable enters the objective with a negative coefficient and consumes capacity in (19) and (20). An optimal solution therefore sets each of them to the smallest value permitted by (24)-(28), which is exactly the product it represents once the two binary factors are fixed. The upper-bounding constraints of the standard linearization, namely $\tilde{x}_{k,a}(\xi) \leqslant x_{k,a}$, $\tilde{x}_{k,a}(\xi) \leqslant l_k(\xi)$ and their counterparts, are therefore redundant and are omitted, and the auxiliary variables can be declared continuous.

The remaining nonlinear term of (15), $\phi_k(1 - e_k(\xi))z_k$, requires no auxiliary variable. Constraint (18) imposes $e_k(\xi) \leqslant z_k$ with both variables binary, so that $e_k(\xi)z_k = e_k(\xi)$ and the term reduces to $\phi_k(z_k - e_k(\xi))$.

This linearization is required only when the extensive form is solved directly. In the reformulation of Section 5.1, the first-stage variables are fixed, so all these products are linear in $l_k(\xi)$ and no auxiliary variables are needed.

In contrast to classical stochastic service network design models (e.g., Lium et al., 2009; Crainic et al., 2011), where network design decisions are the exclusive province of the first stage and commodity flows are second-stage recourse, the SSND-UDS assigns both service selection and shipment routing to the first stage. The second stage

is restricted to adapting the execution of pre-committed itineraries. This structure is motivated by two features of M1M platforms. First, the IDSP does not own transportation assets; the capacity over which tactical itineraries are built is sourced from carriers through contractual agreements that must be secured before operations begin. Second, the IDSP communicates confirmed shipment itineraries to shippers and carriers well in advance of execution. Consequently, second-stage recourse is structurally constrained to shipment-level execution decisions: following the tactical itinerary, outsourcing to an ad hoc carrier, or incurring a non-fulfillment penalty.

5 Solution approach

Evaluating the expected recourse value in (1) requires integrating over the joint distribution of shipment volumes and service capacities, which is not tractable. We therefore adopt the sample average approximation (Kleywegt et al., 2002; Taherkhani et al., 2025), in which a finite sample $S = \{\xi^1, \ldots, \xi^{|S|}\}$ of realizations is drawn from $\mathbb{P}$ and the expectation is replaced by the corresponding sample average. Each realization is assigned equal probability $p_s = 1/|S|$, so that $\sum_{s \in S} p_s = 1$. This yields the *extensive form*

$$\max \sum_{k \in \mathcal{K}^C} \rho_k - \sum_{\sigma \in \Sigma^N} f_\sigma y_\sigma + \sum_{s \in S} p_s \, Q(x, y, z, r, v; \xi^s), \tag{29}$$

subject to (2)-(14), where $Q(\cdot\,; \xi^s)$ is given by (15)-(23). For readability, we index by s the second-stage variables and the realized parameters associated with ξ^s, writing $w^s_k = w_k(\xi^s)$, $u^s_a = u_a(\xi^s)$, $e^s_k = e_k(\xi^s)$, $l^s_k = l_k(\xi^s)$, and $h^s_{k,\sigma} = h_{k,\sigma}(\xi^s)$. We further alleviate the notation by referring to ξ^s simply as the realization corresponding to scenario s. The distributions from which the sample is drawn, and their calibration by stakeholder category, are described in Section 6.

For the smaller instances, the extensive form, linearized as described in Section 4.3, is solved directly with Gurobi. Solving it quickly becomes impractical, however, as the number of scenarios and the size of the SSND components, e.g., the number of services and shipment requests, increase. We therefore apply the decision-based scenario clustering method of Hewitt et al. (2022), accelerated through multi-threaded processing, to compute valid bounds more efficiently, as described in Section 5.2. This method requires evaluating fixed first-stage decisions over many scenarios. To perform these evaluations efficiently, we first show, in Section 5.1, that for fixed first-stage decisions the second-stage problem can be equivalently reformulated as a multidimensional knapsack problem (MKP).

5.1 MKP reformulation for the scenario subproblems

Once the first-stage variables $(\hat{y}, \hat{z}, \hat{x}, \hat{v}, \hat{r}^o, \hat{r}^d)$ are fixed, the second-stage decisions decouple by scenario. In each scenario s, the recourse problem determines, for each shipment, whether to execute its tactical itinerary or to fall back to an alternative (ad hoc service or non-fulfillment), subject to realized capacities. We show that this recourse admits an equivalent *0–1 multidimensional knapsack* (MKP) reformulation.

Let

$$\mathcal{K}^{SE}(\hat{z}) \;=\; \mathcal{K}^C \;\cup\; \big\{\, k \in \mathcal{K}^N \cup \mathcal{K}^I \;:\; \hat{z}_k = 1 \big\} \tag{30}$$

denote the set of *selected* shipments, i.e., all contract-based shipments, together with the non-contract and irregular shipments accepted in the first stage. For notational brevity, we suppress the dependence on $\hat{z}$ and write $\mathcal{K}^{SE}$ throughout.

For each $k \in \mathcal{K}^{SE}$ and $a \in \mathcal{A}$, define

$$\pi_k(s) \;=\; \begin{cases} -\displaystyle\sum_{t \in \mathscr{T}(o_k)} \psi^P_{k,t}\, \hat{r}^o_{k,t} - \sum_{t \in \mathscr{T}(d_k)} \psi^D_{k,t}\, \hat{r}^d_{k,t} - \sum_{a \in \mathcal{A}^M} c^k_a\, w^s_k\, \hat{x}_{k,a} - \sum_{a \in \mathcal{A}^H} c^k_a\, w^s_k\, \hat{v}_{k,a}, & k \in \mathcal{K}^C, \\ \rho_k - \displaystyle\sum_{t \in \mathscr{T}(o_k)} \psi^P_{k,t}\, \hat{r}^o_{k,t} - \sum_{t \in \mathscr{T}(d_k)} \psi^D_{k,t}\, \hat{r}^d_{k,t} - \sum_{a \in \mathcal{A}^M} c^k_a\, w^s_k\, \hat{x}_{k,a} - \sum_{a \in \mathcal{A}^H} c^k_a\, w^s_k\, \hat{v}_{k,a}, & k \in \mathcal{K}^N \cup \mathcal{K}^I, \end{cases} \tag{31}$$

$$b_k(s) \;=\; \begin{cases} -\displaystyle\min_{\sigma \in \Sigma^A(k)} c^k_\sigma\, w^s_k, & k \in \mathcal{K}^C, \\ \max\Big\{ \rho_k - \displaystyle\min_{\sigma \in \Sigma^A(k)} c^k_\sigma\, w^s_k,\; -\phi_k \Big\}, & k \in \mathcal{K}^N \cup \mathcal{K}^I, \end{cases} \tag{32}$$

$$\zeta_{k,a}(s) \;=\; w^s_k \big(\hat{x}_{k,a} + \hat{v}_{k,a} \big), \tag{33}$$

$$\gamma_a(s) \;=\; \begin{cases} u^s_a\, \hat{y}_{\sigma(a)}, & \text{if } \sigma(a) \in \Sigma^N, \\ u^s_a, & \text{otherwise}, \end{cases} \tag{34}$$

where $\pi_k(s)$ is the net profit when shipment k follows its tactical itinerary, $b_k(s)$ is the optimal fallback value, $\zeta_{k,a}(s)$ is the capacity consumed on arc a, and $\gamma_a(s)$ is the available capacity. Note that ρ_k for contract-based shipments is excluded from π_k because it already appears in the first-stage objective (1).

Proposition 5.1. *For any feasible first-stage solution $(\hat{y}, \hat{z}, \hat{x}, \hat{v}, \hat{r}^o, \hat{r}^d)$ and any scenario $s \in \mathcal{S}$, the recourse value $Q(\hat{y}, \hat{z}, \hat{x}, \hat{v}, \hat{r}^o, \hat{r}^d;\, s)$ defined in (15) satisfies*

$$Q(\hat{y}, \hat{z}, \hat{x}, \hat{v}, \hat{r}^o, \hat{r}^d;\, s) \;=\; \sum_{k \in \mathcal{K}^{SE}} b_k(s) \;+\; \mathrm{MKP}(s), \tag{35}$$

where

$$\text{MKP}(s) \;=\; \max \quad \sum_{k\in\mathcal{K}^{SE}} \big(\pi_k(s) - b_k(s)\big)\, l_k^s \tag{36}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{K}^{SE}} \zeta_{k,a}(s)\, l_k^s \;\leqslant\; \gamma_a(s), \qquad \forall\, a \in \mathcal{A}, \tag{37}$$

$$l_k^s \in \{0,1\}, \qquad \forall\, k \in \mathcal{K}^{SE}, \tag{38}$$

Proof. Fix $s \in \mathcal{S}$. The second-stage variables for each shipment k are the itinerary-use indicator l_k^s, the satisfaction indicator e_k^s (for $k \in \mathcal{K}^N \cup \mathcal{K}^I$), and the ad hoc assignment $h_{k,\sigma}^s$. We eliminate e_k^s and $h_{k,\sigma}^s$ by case analysis, reducing the recourse to decisions over l_k^s alone.

Case 1: Contract-based shipments ($k \in \mathcal{K}^C$). Constraint (16) requires $\sum_{\sigma\in\Sigma^A(k)} h_{k,\sigma}^s + l_k^s = 1$, so the shipment either follows its tactical itinerary ($l_k^s = 1$, $h_{k,\sigma}^s = 0$ for all σ) or is outsourced to exactly one ad hoc service ($l_k^s = 0$). If $l_k^s = 1$, the contribution to (15) from shipment k is $\pi_k(s)$: the penalty costs for the fixed pickup and delivery times, plus the transportation and holding costs along its itinerary. If $l_k^s = 0$, the recourse selects the cheapest ad hoc service, contributing $b_k(s) = -\min_{\sigma\in\Sigma^A(k)} c_\sigma^k\, w_k^s$.

Case 2: Non-contract and irregular shipments ($k \in \mathcal{K}^N \cup \mathcal{K}^I$, $\hat{z}_k = 1$). Constraints (17)–(18) give $\sum_\sigma h_{k,\sigma}^s + l_k^s = e_k^s \leqslant 1$, yielding three sub-cases. If $l_k^s = 1$: then $e_k^s = 1$ and $h_{k,\sigma}^s = 0$ for all σ, contributing $\pi_k(s)$, which includes the revenue ρ_k. If $l_k^s = 0$ and $e_k^s = 1$: the shipment is fulfilled via the cheapest ad hoc service, contributing $\rho_k - \min_\sigma c_\sigma^k\, w_k^s$. If $l_k^s = 0$ and $e_k^s = 0$: the shipment is not fulfilled, incurring cost $-\phi_k$. When $l_k^s = 0$, the recourse optimally selects the better of the latter two options, so the fallback value is $b_k(s) = \max\{\rho_k - \min_\sigma c_\sigma^k\, w_k^s,\; -\phi_k\}$.

Aggregation. In both cases, shipment k's contribution equals $b_k(s) + (\pi_k(s) - b_k(s))\, l_k^s$, since this reduces to $\pi_k(s)$ when $l_k^s = 1$ and to $b_k(s)$ when $l_k^s = 0$. Summing over all $k \in \mathcal{K}^{SE}$ separates into a constant $\sum_k b_k(s)$ and an optimization over l_k^s. The only coupling across shipments arises from the capacity constraints (19)–(20). Substituting the fixed first-stage assignments, the aggregate volume on arc a becomes $\sum_{k\in\mathcal{K}^{SE}} w_k^s(\hat{x}_{k,a} + \hat{v}_{k,a})\, l_k^s = \sum_{k\in\mathcal{K}^{SE}} \zeta_{k,a}(s)\, l_k^s$, which must not exceed $\gamma_a(s)$. This yields the MKP formulation (36)–(38). □

The decomposition (35) has a natural interpretation: $\sum_k b_k(s)$ is the recourse profit when every shipment uses its best fallback, and the MKP captures the incremental benefit of activating tactical itineraries, where each coefficient $\pi_k(s) - b_k(s)$ measures the value of the itinerary for shipment k over its fallback.

Substituting (35) into (1) yields

$$\max \sum_{k\in\mathcal{K}^C} \rho_k - \sum_{\sigma\in\Sigma^N} f_\sigma\, \hat{y}_\sigma + \sum_{s\in\mathcal{S}} p_s \Big[\sum_{k\in\mathcal{K}^{SE}} b_k(s) + \mathrm{MKP}(s) \Big]. \tag{39}$$

5.2 Scenario clustering derived bounds

The scenario-clustering approach proposed by Hewitt et al. (2022) was introduced as a means of interpreting the effects of uncertainty, represented through a set of possible scenarios, from the perspective of the underlying decision problem. More specifically, this approach enables scenarios to be grouped according to their similarity or dissimilarity within the decision space. It was further shown that, when solving challenging stochastic optimization models, the resulting clusters can improve computational efficiency and yield stronger bounds. In this section, we review the main elements of the approach and explain how it can be applied in the context considered here.

The fundamental idea underlying the approach is to define an *opportunity cost distance* between pairs of scenarios, say (s_1, s_2), by evaluating how well the scenario-specific solution associated with s_1 performs when scenario s_2 occurs. More specifically, let

$$\begin{aligned}(\hat{x}^{s_1}, \hat{y}^{s_1}, \hat{z}^{s_1}, \hat{r}^{s_1}, \hat{v}^{s_1}) = \underset{(x,y,z,r,v)}{\arg\max} \Big\{ \sum_{k\in\mathcal{K}^C} \rho_k - \sum_{\sigma\in\Sigma^N} f_\sigma\, y_\sigma + \; Q(x, y, z, r, v; s_1) \Big\} \\ s.t. \qquad (2) - (23)\end{aligned}$$

be the optimal first-stage decision under scenario s_1. In the present context, the solution $(\hat{x}^{s_1}, \hat{y}^{s_1}, \hat{z}^{s_1}, \hat{r}^{s_1}, \hat{v}^{s_1})$ can be viewed as a signal of how scenario s_1 influences the decisions made when it is predicted to occur. We now denote the *ex post profit* obtained by implementing the scenario-specific solution associated with s_1 when scenario s_2 occurs as

$$\Pi\big(s_1; s_2\big) = \sum_{k\in\mathcal{K}^C} \rho_k - \sum_{\sigma\in\Sigma^N} f_\sigma\, \hat{y}^{s_1}_\sigma \; + \; Q\big(\hat{x}^{s_1}, \hat{y}^{s_1}, \hat{z}^{s_1}, \hat{r}^{s_1}, \hat{v}^{s_1}; s_2\big),$$

where $Q(\cdot; s_2)$ is defined in (15) and also can be calculated by $Q_{MKP}(s)$ from Proposition 5.1. Then the unidirectional opportunity cost is defined as

$$\Delta(s_1, s_2) = \Pi\big(s_2; s_2\big) \; - \; \Pi\big(s_1; s_2\big) \; \geqslant \; 0,$$

i.e., the loss incurred by using the plan for s_1 when scenario s_2 occurs. Finally, we symmetrize the measure to obtain the scenario-distance metric $d(s_1, s_2) = \Delta(s_1, s_2) + \Delta(s_2, s_1)$. Thus, $d(s_1, s_2)$ assesses the dissimilarity between scenarios s_1 and s_2 based on the gaps in objective value that arise when the solution developed in anticipation of either scenario is evaluated under the other.

By considering all pairs of scenarios in S and computing the corresponding scenario distance for each pair, we obtain a distance matrix $\boldsymbol{D}$ of dimension $|S| \times |S|$. The clustering approach then uses this matrix to partition the scenario set S into κ non-overlapping clusters. In the present context, these partitions are constructed to derive valid bounds for the extensive model. More specifically, the aim is to determine how scenarios should be grouped so that they can be effectively used within the different bounding techniques employed throughout the optimization process.

The characteristics of each partition depend on the bounding technique for which it is constructed. In particular, clusters may be formed by either minimizing or maximizing intra-cluster distances, depending on whether the objective is to group scenarios with lower or higher levels of decision-based dissimilarity. Moreover, we do not impose additional structural constraints on the clusters themselves. We only specify the number of clusters, denoted by κ, that defines the partition.

To obtain partitions with minimal intra-cluster distance, we apply the Partitioning Around Medoids (PAM) algorithm, also known as k-medoids (Kaufman and Rousseeuw, 1990). The objective of PAM is to partition the observations into κ clusters such that each cluster is represented by a *medoid*, chosen from its own observations, and the total dissimilarity between observations and their cluster medoids is minimized. It differs from the k-means algorithm (Hartigan and Wong, 1979) by the fact that it only requires a distance matrix D instead of a Euclidean representation and by guaranteeing that the representatives are elements of S itself. The medoids are selected greedily, beginning with the scenario whose total distance to all others is smallest and adding at each step the scenario that most reduces the total deviation. This initial set is then improved by repeatedly exchanging a medoid with the non-medoid scenario that most reduces the total deviation, until no exchange yields an improvement, and both steps follow deterministic rules, so the partition obtained depends only on D and is reproducible across runs. In this way, a partition $\breve{\mathcal{C}} = \{\breve{C}_1, \ldots, \breve{C}_\kappa\}$ that provides maximally homogeneous elements (with respect to D) is obtained.

On the other hand, we can also obtain a partition with maximally heterogeneous elements. This is known as *inverse clustering*. While many inverse clustering algorithms exist in the literature, we propose a simple approach: redistribute the scenarios so that each cluster of the inverse partition contains observations from each of $\breve{C}_1, ..., \breve{C}_\kappa$. In other words, the inverse clustering interleaves the clusters from $\breve{C}$, ensuring that every inverse cluster is composed of observations drawn from all original clusters. Denote the partition obtained in this way by $\widehat{\mathcal{C}} = \{\widehat{C}_1, \ldots, \widehat{C}_\kappa\}$.

We apply the three bounding techniques originally developed in Hewitt et al. (2022). The first two are the *Cluster Bounds*, namely an upper-bounding technique and a lower-bounding technique. The third, referred to as the *Medoid Bound*, defines a lower-bounding technique in the present context. These techniques are described below. To

simplify the notation, we describe them in terms of a partition $C = \{C_1, \ldots, C_k\}$, which can be either the PAM partition $\breve{C}$ or the inverse partition $\widehat{C}$.

For a first-stage decision $\chi = (x, y, z, r, v)$ feasible for (2)–(14), we extend the notation introduced above and write $\Pi(\chi; s)$ for the ex post profit obtained when χ is implemented and scenario s occurs, so that $\Pi(s_1; s_2) = \Pi(\hat{\chi}^{s_1}; s_2)$. The expected profit of χ over the complete scenario set is

$$Z(\chi) = \sum_{s \in \mathcal{S}} p_s \, \Pi(\chi; s) = \sum_{k \in \mathcal{K}^C} \rho_k - \sum_{\sigma \in \Sigma^N} f_\sigma \, y_\sigma + \sum_{s \in \mathcal{S}} p_s \, Q_{\mathrm{MKP}}(\chi; s),$$

obtained by fixing χ and solving the MKP reformulation (Proposition 5.1) for each scenario. Since χ is feasible for the extensive model, $Z(\chi) \leqslant Z^*$, where Z^* denotes its optimal value. Given a cluster C_j of the partition, we write $p_{C_j} = \sum_{s \in C_j} p_s$ and denote by M_j the restricted model obtained by limiting the scenario set to C_j and normalizing the scenario probabilities within the cluster. Its optimal first-stage decision is denoted $\chi^{(j)}$.

Cluster Bounds. Solving M_j for every $j = 1, \ldots, \kappa$ relaxes the non-anticipativity of the first-stage decisions across clusters. Each cluster then receives its own plan, as if the cluster of the realized scenario were known in advance. Aggregating the cluster optima with the cluster probabilities gives the *cluster upper bound*

$$UB_{\mathrm{cluster}} = \sum_{j=1}^{\kappa} p_{C_j} \, \mathrm{OptVal}(M_j). \tag{40}$$

This relaxation enlarges the feasible decision space of a maximization problem, so that $UB_{\mathrm{cluster}} \geqslant Z^*$.

The same partition also yields κ candidate decisions $\chi^{(1)}, \ldots, \chi^{(\kappa)}$, each optimal for the scenarios of its own cluster. Every candidate is feasible for the extensive model, since the first-stage constraints do not depend on the scenario set. Each cluster is thus treated as a sample of $\mathcal{S}$, and the candidate it produces is evaluated against the complete scenario set. Retaining the best of them gives the *cluster lower bound*

$$LB_{\mathrm{cluster}} = \max_{1 \leqslant j \leqslant \kappa} Z(\chi^{(j)}) \;\leqslant\; Z^*. \tag{41}$$

Medoid Lower Bound. Again following Hewitt et al. (2022), a second lower bound is obtained from the medoids of the clusters. The medoid σ_j of cluster C_j is the scenario with the smallest average distance to the other scenarios of that cluster, and it is assigned the total probability p_{C_j} of its cluster. Solving the reduced model that retains only the medoid scenarios $\sigma_1, \ldots, \sigma_\kappa$ with these probabilities yields a candidate first-stage

decision χ^{med}. The reduction keeps the size of the model manageable while limiting the information lost, and χ^{med} is again feasible for the extensive model, so that

$$LB_{\text{medoid}} = Z(\chi^{\text{med}}) \leqslant Z^*. \tag{42}$$

The three techniques are valid for any partition of $\mathcal{S}$. The partition retained affects the tightness of the resulting bounds rather than their validity, and the gap between LB_{cluster}, LB_{medoid} and UB_{cluster} certifies the quality of the candidate plans when the extensive model cannot be solved directly.

Multi-thread processing for distance matrix calculation. Computing the distance matrix requires $|\mathcal{S}|$ single–scenario solves and $2\binom{|\mathcal{S}|}{2}$ MKP optimizations. The computation is parallel at two levels. First, the $|\mathcal{S}|$ single–scenario solves that yield $\hat{x}^s$ are independent. Second, for each unordered pair $\{s_1, s_2\}$, one worker computes the two directed recourse values $Q_{\text{MKP}}(\hat{x}^{s_1}; s_2)$ and $Q_{\text{MKP}}(\hat{x}^{s_2}; s_1)$ in one batch. In the second step, to reduce per-task latency, we use the MKP transformation instead of the main model.

6 Experimental Results and Analysis

This section reports the results of a comprehensive computational study designed to address four objectives. First, we assess the computational tractability of the SSND-UDS formulation when solved directly with Gurobi, and evaluate the quality of the scenario clustering bounds for instances where direct solution is impractical. Second, we examine the impact of the sample size on the stochastic model through in-sample and out-of-sample experiments, which also determine the number of scenarios retained for the remaining experiments. Third, we quantify the economic value of stochastic planning over deterministic planning through the Value of the Stochastic Solution (VSS). Fourth, we analyze the operational properties of stochastic plans in terms of tactical itinerary adherence and recourse cost efficiency.

6.1 Experimental Design

We construct benchmark instances by varying the three parameters of the network structure shown in Table 4. Each configuration is replicated three times with different random seeds, yielding 54 base instances. In all instances, the schedule length spans one week, discretized into $|\mathscr{T}| = 7$ periods of one day each. This granularity captures the dynamics of freight transportation over a typical operational week, while keeping the extensive form

within a size range for which the direct solve can serve as a benchmark on the smaller configurations. Shipment volumes and service capacities are sampled from normal distributions parametrized by a coefficient of variation. For each shipment request $k \in \mathcal{K}$, the volume under scenario s is sampled as

$$w_k^s \sim \mathcal{N}\left(\bar{w}_k, (\nu_k \cdot \bar{w}_k)^2\right), \tag{43}$$

where $\bar{w}_k$ denotes the announced volume for contract-based shippers or the forecasted volume for non-contract and irregular shippers, and $\nu_k > 0$ is the coefficient of variation reflecting demand variability. Similarly, for each service $\sigma \in \Sigma^C \cup \Sigma^N$, the capacity under scenario s is sampled as

$$u_\sigma^s \sim \mathcal{N}\left(\bar{u}_\sigma, (\nu_\sigma \cdot \bar{u}_\sigma)^2\right), \tag{44}$$

where $\bar{u}_\sigma$ is the announced capacity and $\nu_\sigma > 0$ captures capacity variability. To ensure physical validity, volume and capacity realizations are truncated at zero. Scenarios are generated by Monte Carlo sampling, following standard practice in sample average approximation implementations (Kleywegt et al., 2002; Taherkhani et al., 2025): for each scenario $s = 1, \ldots, |S|$, all uncertain parameters are sampled independently from their respective distributions.

Table 4: Network structure parameters

Parameter	**Symbol**	**Values**
Terminals	$\lvert\mathcal{N}^{PH}\rvert$	10, 20, 30
Periods	$\lvert\mathcal{T}\rvert$	7
Shipments	$\lvert\mathcal{K}\rvert$	50, 80, 100
Services per terminal pair	m	2, 4
Replications	–	3
Total base instances: $3 \times 3 \times 2 \times 3 = 54$		

The choice of normal distributions follows common practice in the stochastic service network design literature (Lium et al., 2009; Hewitt et al., 2019; Taherkhani et al., 2025) and provides a parsimonious representation of business-as-usual variability around announced or forecasted values. Contract-based and non-contract carriers operate their services over repeated seasons, so that realized capacities can be viewed as long-run outcomes of multiple operational fluctuations. When these fluctuations arise from the aggregation of several independent or weakly dependent factors, a normal distribution approximates the variability around the expected capacity with few parameters. The sample average approximation framework itself imposes no restriction on the form of the distributions from which scenarios are drawn, and alternative choices could be adopted without modifying the formulation or the solution approach.

Reliability differences across stakeholder categories are represented through the coefficient of variation rather than through the mean. Since the standard deviation of a

service capacity is set proportionally to the announced capacity, i.e., equal to $\nu_\sigma \bar{u}_\sigma$, a shortfall of a given percentage has the same probability for every service in a category, regardless of the capacity announced. A service offering more capacity is therefore not more reliable. The coefficient of variation is differentiated by category to reflect the reliability hierarchy inherent in the M1M system, grounded in the operational characteristics of each stakeholder type described in Section 3.

On the demand side, contract-based shippers ($k \in \mathcal{K}^C$) have established strategic partnerships with the platform and negotiate capacity requirements in advance. This relationship provides the IDSP with better demand visibility, though volumes may still fluctuate due to production delays or market variations. We set $\nu_k \in [0.15, 0.25]$ for $k \in \mathcal{K}^C$. Regular shipment requests from non-contract shippers ($k \in \mathcal{K}^N$) are submitted well in advance of their release dates, providing some predictability, but the absence of binding commitments introduces additional variability. We set $\nu_k \in [0.25, 0.40]$ for $k \in \mathcal{K}^N$. Irregular shipment requests ($k \in \mathcal{K}^I$) are forecasted from historical data and are inherently speculative, being less predictable and supplementary rather than central to the platform's operations. We set $\nu_k \in [0.40, 0.60]$ for $k \in \mathcal{K}^I$.

On the supply side, contract-based carriers ($\sigma \in \Sigma^C$) operate under long-term agreements and represent the most reliable capacity source. Nevertheless, actual capacities may fall short of agreed terms due to equipment failures, driver illness, traffic congestion, or strikes. We set $\nu_\sigma \in [0.10, 0.20]$ for $\sigma \in \Sigma^C$. Non-contract carriers ($\sigma \in \Sigma^N$) lack long-term commitments to the platform and prioritize their own profitability, which introduces uncertainty stemming from both unpredictable operational events and carriers' decisions to modify or withdraw service offers. We set $\nu_\sigma \in [0.20, 0.35]$ for $\sigma \in \Sigma^N$. Table 5 summarizes the coefficient of variation ranges by stakeholder category.

Table 5: Coefficient of variation (ν) calibration by stakeholder category

Side	**Category**	**Symbol**	**CoV Range**
Demand	Contract-based shippers	$\mathcal{K}^C$	$[0.15, 0.25]$
	Non-contract shippers	$\mathcal{K}^N$	$[0.25, 0.40]$
	Irregular shippers	$\mathcal{K}^I$	$[0.40, 0.60]$
Supply	Contract-based carriers	Σ^C	$[0.10, 0.20]$
	Non-contract carriers	Σ^N	$[0.20, 0.35]$

For stochastic experiments, we generate $|S| = 100$ scenarios per instance with equal probabilities under three uncertainty regimes: *low*, *medium*, and *high*, corresponding to the lower bounds, midpoints, and upper bounds of the coefficient of variation ranges defined in Table 5, respectively. This sample size results from the in-sample and out-of-sample analysis of Section 6.3, which shows that 100 scenarios yield stable solutions for the instance set considered, while larger samples increase the computational effort without improving stability. This yields 162 stochastic test instances in total. All experi-

ments were conducted on compute nodes equipped with two Intel Xeon 6972P processors at 2.4 GHz, providing 192 cores and 384 MB of L3 cache per node, with 240 GB of RAM. Each run was allocated eight cores. All models were solved with Gurobi 12.0.0 under a time limit of 24 hours and default settings otherwise.

6.2 Optimality Results

We assess the computational tractability of the SSND-UDS formulation through two complementary lenses: the performance of Gurobi when solving the full extensive-form MIP directly, and the quality of the decision-based scenario clustering bounds when direct solution is impractical. Both analyses share the same benchmark: the direct-solve upper bound reported by Gurobi at termination.

6.2.1 Direct Solution

Table 6 reports the direct-solve time and optimality gap alongside the scenario distance-matrix computation time used by the clustering procedure of Section 5.2, for a fixed scenario count of $|S| = 100$, stratified by (Terminals, Shipments, Services per arc) and aggregated over three replications per configuration.

Table 6: Effect of instance size on the direct (Gurobi) solve and on the scenario distance-matrix computation for all instances with $|\mathcal{S}| = 100$.

Inst.			**Direct Solve**					**Dist. Calc.**	
			Time (s)			**(%)**		**Time (s)**	
Term.	**Ship.**	**Serv.**	**Mean**	**Std**	**Med.**	**Mean Gap**	**Std Gap**	**Mean**	**Std**
10	50	2	530.0	258.5	434.8	0.00	0.00	1058.5	102.0
10	50	4	6321.2	8953.2	643.3	0.00	0.00	1593.3	284.3
10	80	2	46916.7	31042.7	52920.1	0.14	0.43	1519.7	129.7
10	80	4	31110.6	9322.0	32035.4	0.00	0.00	2177.6	181.0
10	100	2	63973.3	28573.1	86224.9	5.55	13.39	1858.3	40.2
10	100	4	86226.1	2.6	86225.2	3.07	3.32	2296.6	200.6
20	50	2	32708.6	29590.2	19716.1	0.29	0.86	4135.9	221.2
20	50	4	73249.0	24950.8	86254.8	4.31	4.11	6544.3	359.5
20	80	2	86276.9	12.7	86271.8	10.64	16.42	5639.9	652.8
20	80	4	86291.9	5.8	86290.2	12.62	7.44	8470.1	249.1
20	100	2	86291.9	7.0	86289.3	12.23	16.00	6252.9	126.4
20	100	4	86309.9	10.1	86313.6	24.82	15.25	9644.2	733.6
30	50	2	19667.1	17691.1	17157.0	0.00	0.00	9648.8	785.2
30	50	4	64316.6	33684.8	86326.9	2.22	2.36	16453.0	1932.4
30	80	2	81484.0	10056.2	86334.7	14.56	21.91	11310.4	933.3
30	80	4	86384.3	10.7	86381.5	3.18	0.90	19179.1	1384.1
30	100	2	86394.9	3.6	86395.0	21.34	6.32	13498.1	701.0

The results reveal a sharp tractability frontier. For the smallest configuration (10 terminals, 50 shipments, two services per arc), Gurobi proves optimality in 530 seconds on average. The mean gap remains below 0.3% for all groups with 10 terminals and at most 80 shipments, and for the groups with 50 shipments and two services per arc. Difficulty then escalates rapidly with problem size. At 10 terminals and 100 shipments, the median solve time reaches the 24-hour limit and the mean gaps are 3.07% and 5.55%. With 20 or 30 terminals and at least 80 shipments, the median solve time reaches the 24-hour limit in all seven groups, and the mean gap exceeds 10% in six of them. In contrast, the distance-matrix computation remains tractable for all groups, growing from approximately 1,000 to 19,000 seconds, well below the 24 hours consumed by an unsuccessful direct solve at comparable sizes. These findings confirm that the extensive form can be solved reliably only for small networks, and motivate the clustering-based bounds for larger instances.

6.2.2 Scenario Clustering Bounds

We evaluate the bounds of Section 5.2 on the PAM partition $\check{C}$ and the inverse partition $\hat{C}$, for $\kappa \in \{10, 20, 30, 40\}$. Each restricted model M_j is solved with a time limit of 1,600 seconds. When this limit is reached, UB_{cluster} uses the best bound reported by Gurobi for M_j, so that it remains valid, and $\chi^{(j)}$ is the best incumbent found. The medoid bound is computed on $\check{C}$ only, since it requires each medoid to represent the scenarios of its cluster, which the inverse partition does not provide by construction.

Table 7 compares the bounds with the optimal value on the 61 instances for which the direct solve proves optimality. The gap is computed as $100 \times (OPT_{\text{Direct}} - \Upsilon)/OPT_{\text{Direct}}$, where Υ is the bound in the first column, so that it is positive for lower bounds and negative for upper bounds.

The inverse partition yields tighter cluster bounds at every value of κ, in both mean and worst gap. Under both partitions, the cluster bounds loosen as κ increases, so that $\kappa = 10$ yields the tightest Cluster-LB and Cluster-UB. The medoid bound follows the opposite trend and is the tightest lower bound for $\kappa \geqslant 30$. Selecting κ per instance by the silhouette score of the PAM partition yields looser cluster bounds than the fixed choice $\kappa = 10$ (Appendix A.1).

Table 7 is limited to the 61 instances solved to optimality, which are among the smallest of the benchmark. For the remaining instances, the quality of the bounds can only be assessed through the certified gap $100 \times (UB_{\text{cluster}} - LB_{\text{cluster}})/UB_{\text{cluster}}$, which requires no optimal value. Table 8 reports this gap over all 162 instances, by value of κ and by uncertainty regime at $\kappa = 10$, together with the time to compute the cluster bounds.

As shown in Table 8, the inverse partition reduces the mean certified gap at every

Table 7: Summary statistics (in %) for clustering-based bounds under the PAM and inverse partitions, on the subset of instances where the direct solve proves optimality ($LB_{\text{Direct}} = UB_{\text{Direct}}$, 61 instances). Worst Gap is the gap of largest absolute value.

Measure	Partition	Mean Gap	Median Gap	Std Gap	Worst Gap
$\kappa = 10$					
Medoid-LB	PAM	4.04	1.82	5.47	34.23
Cluster-LB	PAM	2.03	0.87	2.62	13.14
Cluster-LB	Inverse	0.93	0.41	1.25	5.45
Cluster-UB	PAM	−4.55	−3.78	3.37	−16.18
Cluster-UB	Inverse	−2.19	−1.53	2.16	−9.63
$\kappa = 20$					
Medoid-LB	PAM	2.33	1.01	3.48	21.38
Cluster-LB	PAM	2.32	1.32	2.91	16.46
Cluster-LB	Inverse	1.39	0.60	1.78	7.24
Cluster-UB	PAM	−5.73	−4.80	4.08	−19.47
Cluster-UB	Inverse	−3.68	−2.72	3.34	−15.31
$\kappa = 30$					
Medoid-LB	PAM	1.46	0.75	1.92	10.05
Cluster-LB	PAM	2.61	1.55	3.33	16.46
Cluster-LB	Inverse	1.59	0.72	2.16	11.23
Cluster-UB	PAM	−6.41	−5.40	4.47	−21.10
Cluster-UB	Inverse	−4.56	−3.58	3.90	−16.78
$\kappa = 40$					
Medoid-LB	PAM	1.16	0.66	1.71	8.68
Cluster-LB	PAM	2.78	1.25	3.56	15.38
Cluster-LB	Inverse	1.72	0.73	2.60	13.69
Cluster-UB	PAM	−6.89	−5.85	4.77	−21.97
Cluster-UB	Inverse	−5.22	−4.22	4.29	−18.64

Table 8: Mean certified gap and computation time of the cluster bounds under the PAM and inverse partitions, over all 162 instances (top) and by uncertainty regime at $\kappa = 10$, 54 instances each (bottom). The reduction is expressed in percentage points (pp).

	Certified gap (%)			Time (s)	
	PAM	Inverse	Reduction (pp)	PAM	Inverse
$\kappa = 10$	6.86	3.48	3.38	2,058	1,520
$\kappa = 20$	7.84	5.23	2.61	1,785	1,207
$\kappa = 30$	8.57	6.49	2.08	1,847	1,424
$\kappa = 40$	9.34	7.20	2.14	1,958	1,642
$\kappa = 10$, **by uncertainty regime**					
Low	3.80	1.19	2.60	1,957	1,443
Medium	5.51	2.68	2.83	2,047	1,461
High	11.27	6.56	4.70	2,170	1,655

value of κ, most strongly at $\kappa = 10$, where the gap is nearly halved. At this value of κ, the reduction holds in each uncertainty regime and grows with the level of uncertainty. The inverse partition also requires less computation time than the PAM partition. With $\kappa = 10$, it therefore provides the tightest certification among the configurations tested, at a lower computational cost.

6.3 In-Sample and Out-of-Sample Analysis

This section plays a double role. The in-sample analysis compares the solutions obtained with samples of different sizes and determines the number of scenarios used in the remaining experiments. The out-of-sample analysis evaluates these solutions on an independent set of scenarios, measuring how well first-stage decisions optimized on a finite sample generalize to unseen scenarios. Both analyses are performed on the two-stage stochastic model (1)–(23).

We select 9 base instances with 10 or 20 terminals, 50 to 100 shipments, and 2 or 4 services per arc. For each instance i and sample size $n \in \{50, 100, 150\}$, we draw three independent scenario sets (replications) $\mathcal{S}^{\text{in}}_{i,n,u}$, $u \in \{1, 2, 3\}$, and solve the resulting extensive-form model with Gurobi under a 6-hour time limit, retaining the best solution found when the limit is reached. For each replication (i, n, u) we record two quantities. The first is the in-sample optimum $\widehat{Z}^{\text{in}}_{i,n,u}$, i.e., the optimal value of (1) computed on $\mathcal{S}^{\text{in}}_{i,n,u}$. The second is the out-of-sample (OOS) objective $\widehat{Z}^{\text{out}}_{i,n,u}$, obtained by fixing the optimal first-stage decisions from (i, n, u) and re-optimizing only the recourse for each scenario in an independent held-out set $\mathcal{S}^{\text{out}}$ of 1,000 equally weighted scenarios drawn i.i.d. from the same distribution.

Table 9 reports, for each instance i and cardinality n, the mean in-sample and out-of-sample objectives ($\overline{Z}^{\text{in}}_{i,n}$ and $\overline{Z}^{\text{out}}_{i,n}$), the relative in–out bias $\Delta_{i,n,u} = |\widehat{Z}^{\text{in}}_{i,n,u} - \widehat{Z}^{\text{out}}_{i,n,u}|/\widehat{Z}^{\text{in}}_{i,n,u}$, reported as mean $\pm$ standard deviation in percentage points; and the replication-level dispersion sd_rep$[\widehat{Z}^{\text{out}}]$.

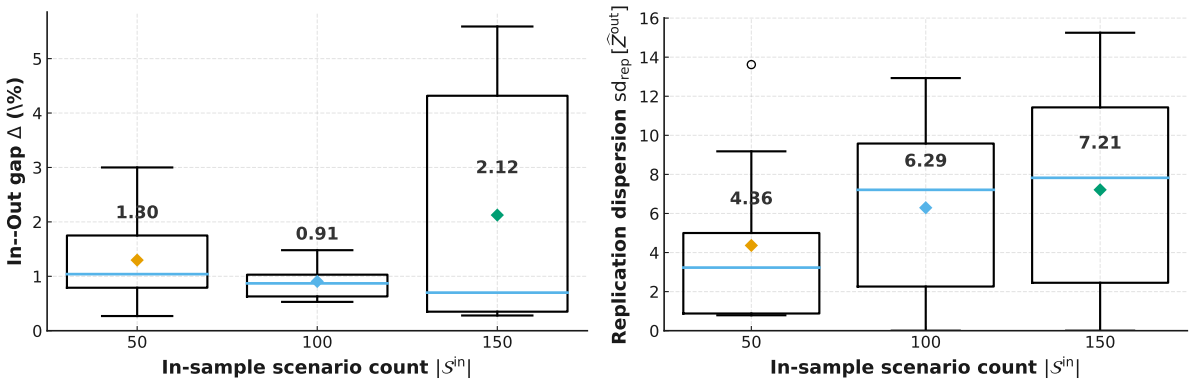
Table 9: Out-of-sample (OOS) stability by instance and in-sample scenario-set size. $\overline{Z}^{\text{in}}$ and $\overline{Z}^{\text{out}}$ are means across replications. $\text{sd}_{\text{rep}}[\widehat{Z}^{\text{out}}]$ is the replication-level dispersion of the OOS mean; $\text{sd}_{\xi}[Z^{\text{out}}]$ is the per-scenario dispersion of the OOS objective over the $|\mathcal{S}^{\text{out}}|$ scenarios. Δ is the relative In–Out gap $|\widehat{Z}^{\text{in}} - \widehat{Z}^{\text{out}}|/\widehat{Z}^{\text{in}}$, reported as $\mu \pm \text{sd}_{\text{rep}}$ in percentage points.

Instance	$\lvert\mathcal{S}^{\text{in}}\rvert$	$\lvert\mathcal{S}^{\text{out}}\rvert$	$\overline{Z}^{\text{in}}$	$\overline{Z}^{\text{out}}$	$\text{sd}_{\text{rep}}[\widehat{Z}^{\text{out}}]$	$\text{sd}_{\xi}[Z^{\text{out}}]$	Δ (%)
p7_t10_k100_m2	50	1000	7,433.8	7,359.3	1.96	641.5	1.04 ± 0.94
p7_t10_k100_m2	100	1000	6,914.1	6,930.6	9.58	730.2	0.97 ± 0.47
p7_t10_k100_m2	150	1000	6,600.7	6,666.7	10.99	682.3	1.02 ± 1.24
p7_t10_k100_m4	50	1000	7,147.7	6,932.8	5.00	949.0	3.00 ± 0.81
p7_t10_k100_m4	100	1000	6,898.0	6,852.6	4.69	985.8	0.86 ± 1.15
p7_t10_k100_m4	150	1000	6,487.8	6,469.1	15.25	914.5	0.32 ± 0.46
p7_t10_k50_m2	50	1000	4,267.1	4,295.8	3.23	370.2	0.80 ± 0.57
p7_t10_k50_m2	100	1000	4,326.8	4,303.0	2.26	367.6	0.55 ± 0.36
p7_t10_k50_m2	150	1000	4,309.5	4,299.8	3.27	359.3	0.36 ± 0.45
p7_t10_k50_m4	50	1000	2,905.5	2,900.7	0.79	321.9	0.48 ± 0.34
p7_t10_k50_m4	100	1000	2,892.1	2,901.0	2.34	320.9	0.63 ± 0.68
p7_t10_k50_m4	150	1000	2,896.2	2,901.0	2.34	320.9	0.38 ± 0.19
p7_t10_k80_m2	50	1000	5,689.8	5,590.2	9.18	656.6	1.75 ± 0.71
p7_t10_k80_m2	100	1000	5,400.1	5,341.2	12.93	648.2	1.23 ± 1.02
p7_t10_k80_m2	150	1000	4,794.0	4,920.6	4.66	926.0	4.94 ± 6.08
p7_t10_k80_m4	50	1000	6,270.9	6,264.1	3.76	444.3	0.27 ± 0.28
p7_t10_k80_m4	100	1000	6,246.9	6,182.4	7.21	484.4	1.03 ± 0.30
p7_t10_k80_m4	150	1000	5,448.1	5,742.8	12.42	499.1	5.59 ± 4.89
p7_t20_k100_m2	50	1000	6,802.4	6,798.0	13.62	389.9	1.33 ± 0.76
p7_t20_k100_m2	100	1000	6,704.1	6,752.6	11.34	400.5	0.87 ± 0.58
p7_t20_k100_m2	150	1000	6,503.4	6,485.1	0.00	399.2	0.28 ± 0.00
p7_t20_k50_m2	50	1000	2,252.6	2,225.9	0.85	325.3	2.22 ± 1.01
p7_t20_k50_m2	100	1000	2,253.8	2,226.3	0.00	323.9	1.48 ± 1.49
p7_t20_k50_m4	50	1000	3,904.5	3,885.3	0.88	371.7	0.79 ± 0.96
p7_t20_k50_m4	100	1000	3,784.1	3,798.5	8.63	372.7	0.53 ± 0.45
p7_t20_k50_m4	150	1000	3,317.5	3,449.8	11.10	377.1	4.11 ± 3.03

Figure 3a summarizes the distribution of in-out gaps and replication dispersion across in-sample sizes. The median gap and its interquartile range decrease substantially when increasing n from 50 to 100; from $n = 100$ to $n = 150$, the mean gap changes little while variability increases considerably. The replication dispersion follows a similar pattern, remaining higher at $n \in 100, 150$ than at $n = 50$, and the difference between $n = 100$ and $n = 150$ is modest. On the basis of these results, we set $|S| = 100$ as the default scenario count in subsequent experiments, as this cardinality yields stable out-of-sample performance without incurring the additional computational burden of larger scenario sets.

6.4 Value of the Stochastic Solution (VSS)

We quantify the economic value of explicitly modeling demand and capacity uncertainty through the Value of the Stochastic Solution (VSS), a standard metric in stochastic programming that measures the expected performance loss when uncertainty collapses to its mean in the planning stage (Birge and Louveaux, 1997). Let RP denote the optimal



(a) In–Out gaps $\Delta_{i,n,\cdot}$ by $n := |\mathcal{S}^{\text{in}}|$. (b) Replication dispersion $\text{sd}_{\text{rep}}[\widehat{Z}^{\text{out}}]$ by n.

Figure 3: Out-of-sample stability across in-sample sizes $n \in \{50, 100, 150\}$. Left: In–Out gap $\Delta_{i,n,\cdot}$; right: replication-level dispersion $\text{sd}_{\text{rep}}[\widehat{Z}^{\text{out}}]$.

objective value of the stochastic (recourse) model. Let $\hat{x}_{EV}$ be the first-stage solution obtained from the *expected-value* (EV) problem, in which all uncertain parameters are replaced by their expectations. Fixing $\hat{x}_{EV}$ and re-optimizing recourse for each scenario yields the *expected result of the EV solution*, $EEV = \sum_{s \in \mathcal{S}} p_s \, Q(\hat{x}_{EV}, s)$, where $Q(\hat{x}_{EV}, s)$ is the optimal second-stage objective value in scenario s given first-stage commitments $\hat{x}_{EV}$. The (absolute) value of the stochastic solution is $VSS = RP - EEV$, and we report its relative magnitude as

$$VSS(\%) = \frac{RP - EEV}{|RP|} \times 100. \tag{45}$$

By construction, $VSS > 0$ indicates that optimizing first-stage decisions against the full scenario distribution increases the expected profit relative to committing to a mean-value plan.

To ensure a like-for-like comparison between RP and EEV, we restrict attention to the 61 instances for which the stochastic model is solved to proven optimality; instance-level results are reported in Appendix B. Across this set, stochastic planning yields economically significant gains on average: the mean VSS is 8.29% (median 4.62%), with substantial dispersion (standard deviation 10.35%; range [0.00%, 52.85%]). A one-sample t-test rejects the null hypothesis of zero mean improvement ($t = 6.26$, $p < 0.001$). The distribution is right-skewed, indicating that while many instances exhibit modest gains, a non-trivial subset experiences large improvements under high uncertainty.

Figure 4 shows a clear monotone increase in VSS with uncertainty level. Mean VSS rises from 3.87% under low uncertainty (CV$\approx$ 15%, $n = 22$) to 7.59% under medium uncertainty (CV$\approx$ 25%, $n = 21$) and 14.53% under high uncertainty (CV$\approx$ 35%, $n = 18$). The regime effect is statistically significant (ANOVA: $F = 6.26$, $p = 0.003$), and VSS is

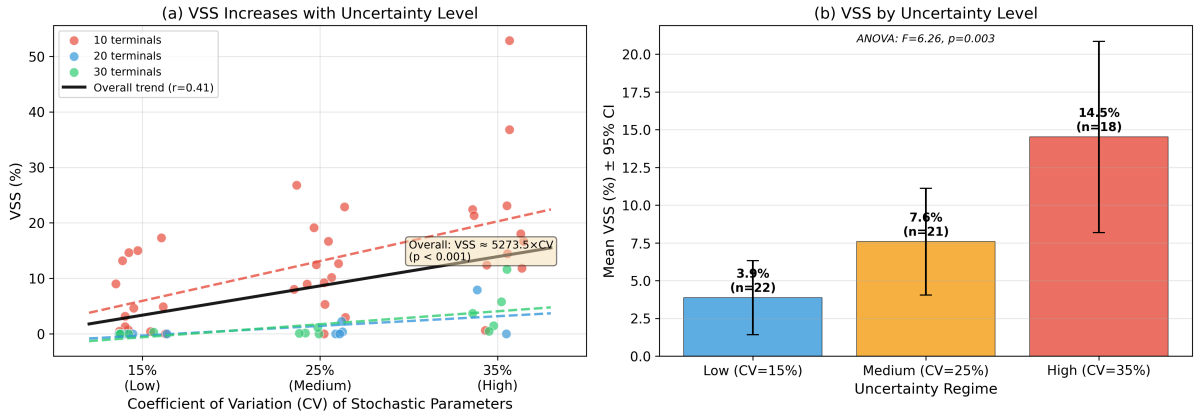


Figure 4: VSS as a function of uncertainty magnitude. Panel (a) shows instance-level VSS versus coefficient of variation (CV), stratified by terminal count; panel (b) reports mean VSS (with 95% CI) by uncertainty regime.

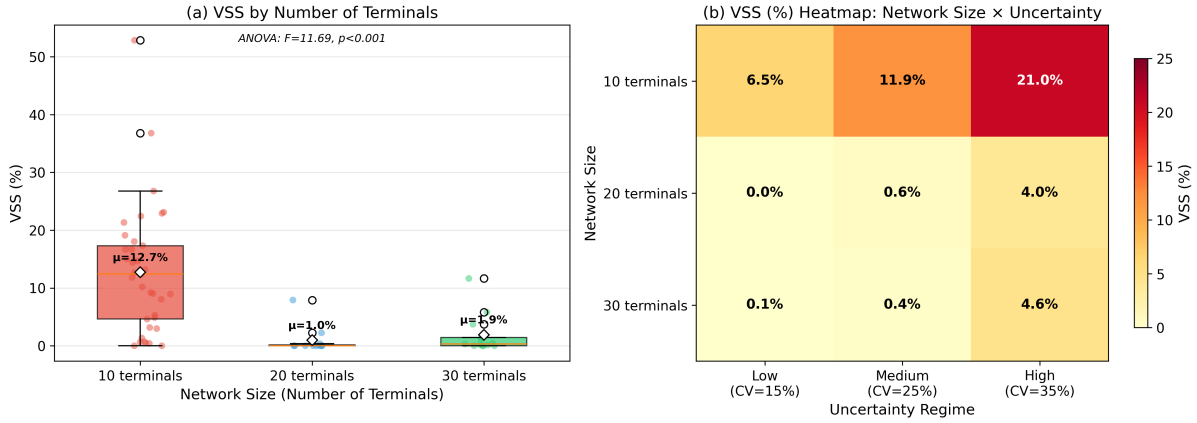


Figure 5: Network-structure effects. Panel (a) shows VSS by terminal count (ANOVA: $F = 11.69$, $p < 0.001$); panel (b) reports the mean VSS heatmap by network size and uncertainty regime.

positively correlated with CV ($r = 0.41$, $p < 0.001$). As forecast errors grow, EEV first-stage allocations increasingly misplace scarce capacity across shipments and itineraries; the stochastic model anticipates adverse realizations and rebalances commitments to protect expected performance.

The network structure moderates VSS substantially. Figure 5 identifies terminal count as the dominant structural determinant: mean VSS is 12.72% for 10-terminal networks ($n = 37$), versus 0.95% for 20 terminals ($n = 11$) and 1.91% for 30 terminals ($n = 13$). The effect is highly significant (ANOVA: $F = 11.69$, $p < 0.001$) and VSS is negatively correlated with terminal count ($r = -0.48$, $p < 0.001$). Larger networks offer more substitution opportunities in recourse, so the EV solution is often near-stochastic in expected performance; in smaller networks, limited substitution makes uncertainty-aware

Table 10: Mean VSS (%) by network size (terminals) and uncertainty regime (CV).

Terminals	**Low (CV≈ 15%)**	**Medium (CV≈ 25%)**	**High (CV≈ 35%)**
10	6.52	11.94	20.95
20	0.00	0.65	3.96
30	0.07	0.36	4.61

hedging materially more valuable.

Within 10-terminal networks, shipment count does not constitute an independent driver of VSS: mean values of 12.57%, 11.88%, and 16.53% for 50, 80, and 100 shipments, respectively, are not statistically distinguishable (one-way ANOVA: $F = 0.27$, $p = 0.763$), and the 100-shipment cell contains only four observations due to computational tractability constraints. We therefore conclude that, within the tractable instance range studied here, shipment count does not constitute an independent driver of VSS; the uncertainty regime is the dominant determinant.

6.5 Decomposing VSS by Uncertainty Source

The aggregate VSS analysis of section 6.4 does not distinguish the relative contributions of demand-side and supply-side uncertainty to the value of stochastic planning. To isolate these effects, we conduct a factorial experiment on 22 base instances spanning 10-, 20-, and 30-terminal networks. For each instance, we solve five configurations: the medium-CV baseline, two variants that change only demand CV (low and high, with supply held at medium), and two variants that change only supply CV (low and high, with demand held at medium), yielding $22 \times 5 = 110$ stochastic programs in total. Table 11 reports the results.

Across all instances and network sizes, supply-side uncertainty contributes substantially more to the VSS than demand-side uncertainty. Increasing supply CV from its low to its high value raises mean VSS by 7.77 percentage points, compared to only 0.52 percentage points for the same change in demand CV. A paired t-test confirms that the supply-side VSS contribution significantly exceeds the demand-side contribution ($t = 3.74$, $p = 0.001$), with $\Delta_s > \Delta_d$ in all 22 instances without exception. The effect is consistent across network sizes: at 10 terminals, the supply-side VSS increment is 10.71 pp versus 0.57 pp for demand; at 20 terminals, 4.51 pp versus 0.28 pp; and at 30 terminals, 2.75 pp versus 0.56 pp.

The sensitivity of the optimal expected profit RP to the source of uncertainty reinforces this finding. Increasing supply CV from its low to its high value reduces RP by 7.11% on average, against 0.94% for the same change in demand CV. In absolute value, the supply effect exceeds the demand effect in 21 of 22 instances (paired t-test, $t = 6.57$,

$p < 0.001$). Supply-side uncertainty is therefore the main driver of the value of stochastic planning in M1M systems, and of the profit loss caused by uncertainty. These results indicate that a tactical plan accounting for demand uncertainty alone would capture only a small part of this value.

Table 11: Decomposing the value of stochastic planning by uncertainty source. Δ_d and Δ_s denote the difference in mean VSS between the high and low CV variants of demand and supply, respectively. RP sensitivity is the relative change in RP, in %, between the high and low CV variants.

	Mean VSS (%)					VSS Δ (pp)		RP Sensitivity (%)	
Terminals	Med.–Med.	Demand CV		Supply CV		Δ_d	Δ_s	Demand	Supply
		Low	High	Low	High				
10	11.53	11.18	11.75	8.62	19.33	+0.57	+10.71	-1.54 ± 1.49	-8.96 ± 3.98
20	0.56	1.05	1.33	0.00	4.51	+0.28	+4.51	-1.07 ± 1.52	-6.64 ± 3.32
30	1.50	0.66	1.22	0.26	3.01	+0.56	+2.75	$+0.71 \pm 0.61$	-2.68 ± 2.51
All	7.26	6.95	7.46	5.16	12.93	+0.52	+7.77	-0.94 ± 1.60	-7.11 ± 4.32

VSS: Paired t-test of $\Delta_s - \Delta_d$: $t = 3.74$, $p = 0.001$; $\Delta_s > \Delta_d$ in 22/22 cases. RP: Paired t-test of |supply| − |demand|: $t = 6.57$, $p < 0.001$; |supply| > |demand| in 21/22 cases.

6.6 Tactical Plan Robustness

Having established the economic value of stochastic planning, we examine whether the first-stage hedging embedded in RP translates into more robust tactical itineraries and lower recourse costs relative to the mean-value plan $\hat{x}_{EV}$. We measure robustness using the *Tactical Itinerary Usage Rate* (TIUR), defined as the fraction of accepted shipments executed on their pre-planned itinerary without rerouting, together with the *Ad-hoc Service Rate* (AHSR) and *Non-fulfillment Rate* (NFR) to characterize the composition of recourse actions.

Aggregating across 61 instances, both solutions achieve high itinerary adherence: TIUR = 94.96% under RP versus 94.34% under $\hat{x}_{EV}$ (Table 12). The improvement is statistically significant (ΔTIUR = +0.62 pp, $p < 0.001$) but modest in magnitude, indicating that the value of stochastic optimization lies not in producing fundamentally different recourse policies but in reducing the cost of recourse interventions.

The economic impact is substantial. As shown in Figure 6(B), RP reduces expected recourse costs by 87.4%, 85.4%, and 82.9% under low, medium, and high uncertainty, respectively, translating to average savings of \$423-\$1,236 per instance. These savings arise almost entirely from reduced ad-hoc service usage (ΔAHSR = −0.62 pp, $p < 0.001$): the deterministic plan remains largely executable, but achieves feasibility by resorting to expensive spot-market capacity that the stochastic plan avoids through better-calibrated

Table 12: Recourse Decision Comparison: Stochastic (RP) vs. Deterministic (EV) Solutions

Category	**Metric**	**RP Mean**	**RP Std**	**EV Mean**	**EV Std**	**Δ (RP-EV)**
			Uncertainty Regime: Low			
Contract	TIUR (%)	95.42	21.31	95.20	21.27	+0.23
	AHSR (%)	0.03	0.05	0.26	0.43	-0.23
Non-contract	TIUR (%)	95.35	21.30	94.78	21.21	+0.56
	AHSR (%)	0.11	0.16	0.67	1.27	-0.56
	NFR (%)	0.00	0.00	0.00	0.00	0.00
Irregular	TIUR (%)	95.40	21.31	94.38	21.17	+1.02
	AHSR (%)	0.06	0.14	1.08	1.97	-1.02
	NFR (%)	0.00	0.00	0.00	0.00	0.00
			Uncertainty Regime: Medium			
Contract	TIUR (%)	95.15	21.80	94.71	21.71	+0.44
	AHSR (%)	0.09	0.10	0.53	0.63	-0.44
Non-contract	TIUR (%)	95.10	21.79	94.03	21.61	+1.07
	AHSR (%)	0.14	0.20	1.21	1.73	-1.07
	NFR (%)	0.00	0.00	0.00	0.00	0.00
Irregular	TIUR (%)	95.10	21.79	93.99	21.59	+1.11
	AHSR (%)	0.14	0.22	1.25	1.54	-1.11
	NFR (%)	0.00	0.00	0.00	0.00	0.00
			Uncertainty Regime: High			
Contract	TIUR (%)	94.22	23.51	93.45	23.34	+0.77
	AHSR (%)	0.23	0.22	0.99	0.90	-0.77
Non-contract	TIUR (%)	94.14	23.50	92.80	23.24	+1.34
	AHSR (%)	0.30	0.33	1.65	1.94	-1.34
	NFR (%)	0.00	0.00	0.00	0.00	0.00
Irregular	TIUR (%)	94.22	23.52	92.73	23.25	+1.49
	AHSR (%)	0.22	0.32	1.71	2.28	-1.49
	NFR (%)	0.00	0.00	0.00	0.00	0.00

Table 13: First-stage decisions under stochastic planning (RP) versus the expected-value plan (EV).

Uncertainty	**Metric**	**RP**	**EV**	**Δ**	**Δ (%)**
Low (CV=0.15)	Services used (Contract)	6.86	6.59	+0.30	+4.1
	Services used (Non-contract)	6.00	5.86	+0.10	+2.3
	Services used (Total)	12.86	12.45	+0.40	+3.3
	Accepted shipments (Contract)	37.64	37.64	0.00	0.0
	Accepted shipments (Non-contract)	10.41	10.86	-0.50	-4.2
	Accepted shipments (Irregular)	4.55	4.82	-0.30	-5.7
Medium (CV=0.25)	Services used (Contract)	7.10	6.62	+0.50	+7.2
	Services used (Non-contract)	5.81	5.81	0.00	0.0
	Services used (Total)	12.90	12.43	+0.50	+3.8
	Accepted shipments (Contract)	36.57	36.57	0.00	0.0
	Accepted shipments (Non-contract)	10.00	10.62	-0.60	-5.8
	Accepted shipments (Irregular)	4.00	4.52	-0.50	-11.6
High (CV=0.35)	Services used (Contract)	6.61	6.11	+0.50	+8.2
	Services used (Non-contract)	5.56	5.50	+0.10	+1.0
	Services used (Total)	12.17	11.61	+0.60	+4.8
	Accepted shipments (Contract)	36.67	36.67	0.00	0.0
	Accepted shipments (Non-contract)	9.56	10.39	-0.80	-8.0
	Accepted shipments (Irregular)	3.89	4.44	-0.60	-12.5

Notes. $\Delta = RP - EV$; $\Delta(\%) = 100 \times \Delta/\text{EV}$.

first-stage decisions. The instance-level scatterplot (Figure 6(C)) confirms a negative relationship between ΔTIUR and ΔCost: instances where RP most improves itinerary adherence are precisely those with the largest recourse cost savings.

Table 13 reveals the underlying mechanism. As uncertainty increases, RP accepts systematically fewer non-contract and irregular shipments than $\hat{x}_{EV}$ (from -4.6% in low to -9.4% in high uncertainty), while contract acceptance remains unchanged, since both solutions are obligated to serve all contract-based shipments. Simultaneously, RP uses more services overall, with increases concentrated in contract-carrier services (from $+4.1\%$ to $+8.2\%$), which provide more reliable capacity and lower recourse risk. The stochastic model therefore hedges not by purchasing excess capacity but by avoiding fragile commitments that the EV plan makes under mean parameters and that later require expensive spot-market corrections.

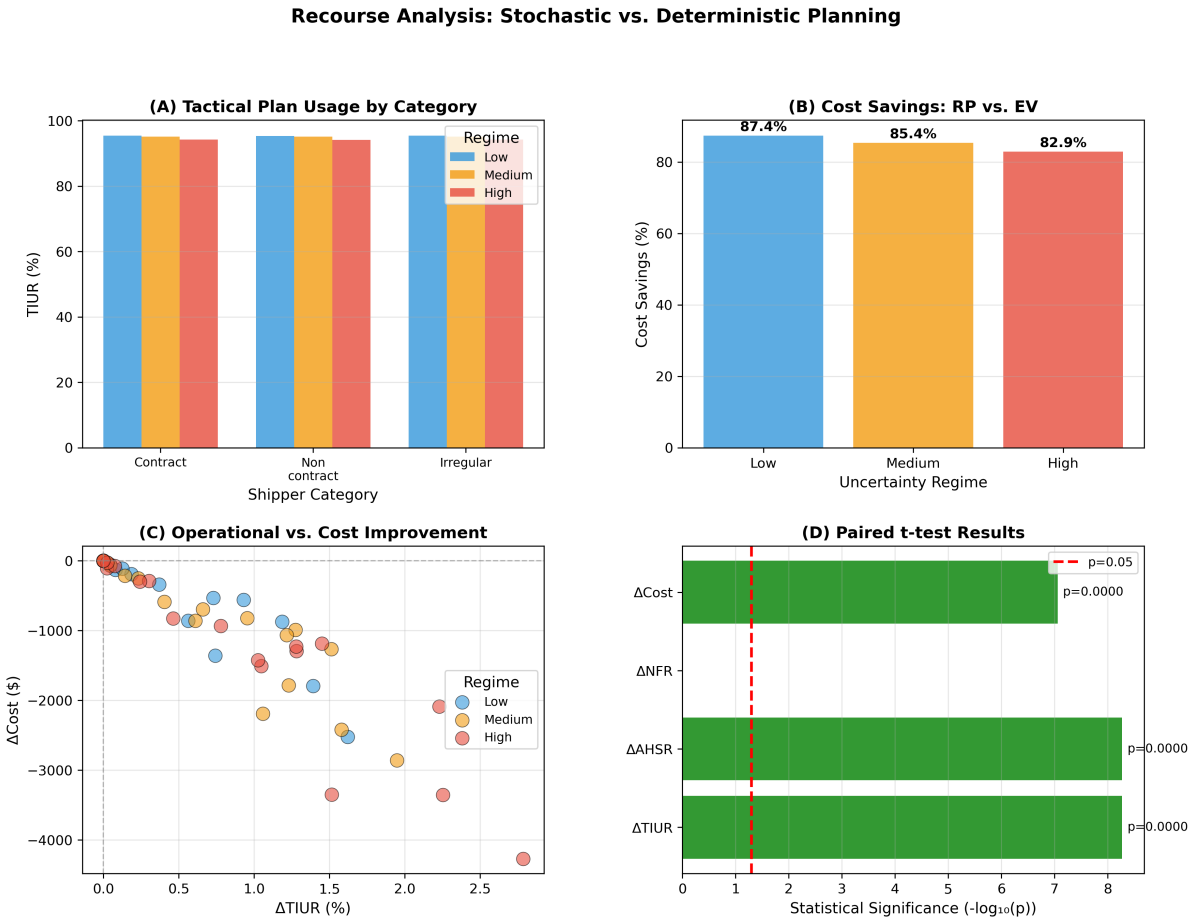


Figure 6: Recourse comparison: stochastic (RP) versus deterministic (EV) planning. (A) Tactical itinerary usage by shipper category and uncertainty regime. (B) Expected recourse cost savings by uncertainty regime. (C) Instance-level relationship between ΔTIUR and ΔCost. (D) Statistical significance of paired differences.

6.7 Non-Fulfillment Penalty Sensitivity Analysis

The non-fulfillment penalty ϕ_k is set by the platform before uncertainty is resolved and is therefore a contract-design parameter rather than a fixed cost. We assess how the platform's plan responds to changes in this parameter by scaling the baseline penalty ϕ_k by a factor $\lambda \in \{1.0, 0.9, 0.75, 0.6, 0.5, 0.4, 0.25, 0.1, 0.05, 0.01\}$ and re-solving the full stochastic model on the 10-terminal instances ($|K| \in \{50, 80\}$, three uncertainty regimes, three replications), with all other parameters and the scenario set held fixed.

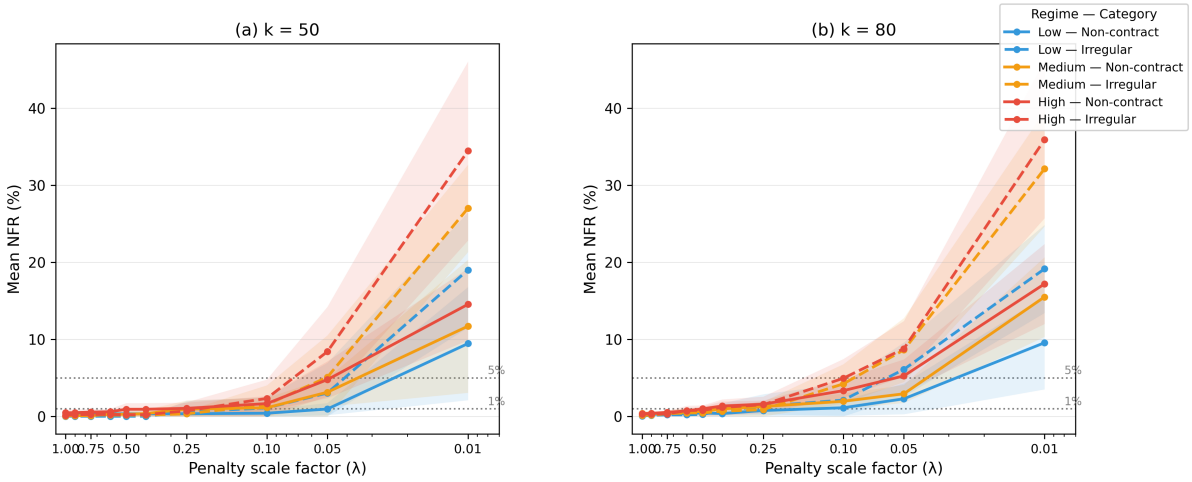


Figure 7: Mean non-fulfillment rate (NFR) as a function of the penalty scale factor λ, by uncertainty regime and shipment category, for (a) $|K| = 50$ and (b) $|K| = 80$ instances. Solid lines: non-contract shipments; dashed lines: irregular shipments. Shaded bands show one standard deviation across replications. Horizontal reference lines mark the 1% and 5% thresholds.

Mean NFR remains at zero across all uncertainty regimes and shipment categories whenever $\lambda \geqslant 0.25$ (Figure 7). Below this threshold, NFR rises gradually until $\lambda = 0.05$ and then sharply at $\lambda = 0.01$, where dropping a shipment becomes cheaper than either executing its itinerary in an overflow scenario or using an ad-hoc service. The threshold is stable across both instance sizes, indicating that it is determined by the relative magnitudes of revenue, ad-hoc costs, and calibrated uncertainty rather than by the shipment mix of a given instance.

At every $\lambda < 0.25$, NFR is higher for irregular shipments than for non-contract shipments. At $\lambda = 0.05$, mean NFR for irregular shipments is 1.7–2.9 times that for non-contract shipments across all uncertainty regimes ($p < 0.001$ in all three regimes; Figure 8). This ordering follows directly from the fulfillment rules in Section 3.2: under a uniform penalty, irregular shipments generate lower expected revenue per unit of capacity and are therefore dropped first when capacity is scarce. The gap between the two categories widens as uncertainty increases, supporting a differentiated penalty structure

with $\phi_k^{\text{irreg}} > \phi_k^{\text{non-contract}}$ to equalize ex post drop rates across categories.

Reducing λ increases expected profit by at most 4–8% relative to the baseline, with the gain flattening beyond roughly 5% aggregate NFR (Figure 9). As λ falls, the platform accepts more non-contract and irregular shipments (Figure 10), but each additional acceptance is more likely to require recourse, which caps the gain. The profit cost of maintaining the baseline penalty is therefore small, while the reputational cost of non-fulfillment is not captured by the objective. Since NFR depends on the prevailing uncertainty regime, the penalty structure should be reviewed and re-calibrated at the start of each planning season to reflect current market conditions.

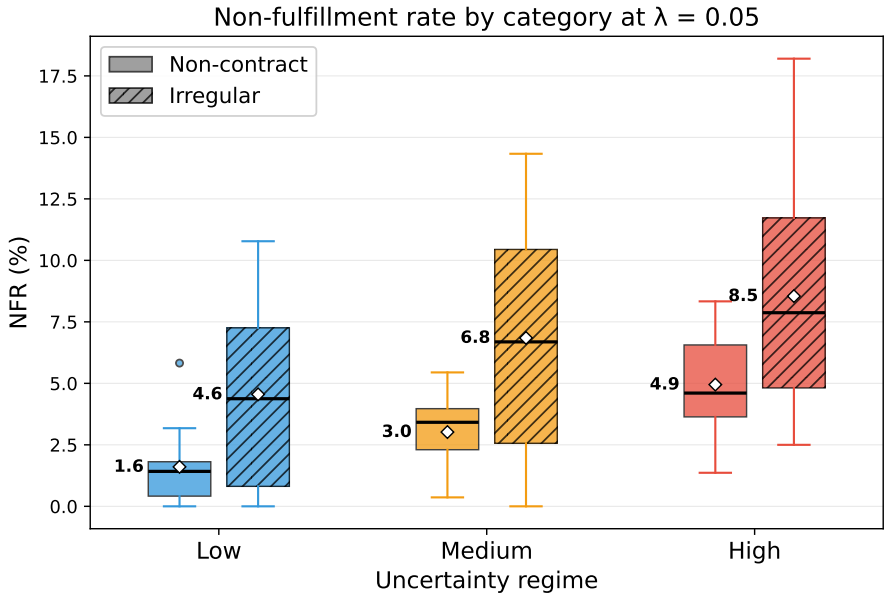


Figure 8: Mean non-fulfillment rate by shipment category and uncertainty regime at $\lambda = 0.05$. Error bars show one standard deviation across instances; the wide intervals reflect instance heterogeneity rather than within-instance noise.



Figure 9: Profit–service trade-off across all (instance, λ) combinations. Each point shows the normalized objective $\mathrm{RP}/\mathrm{RP}_{\lambda=1}$ against the overall non-fulfillment rate; color encodes the penalty scale factor λ. The dashed line marks the $\lambda = 1.0$ baseline.

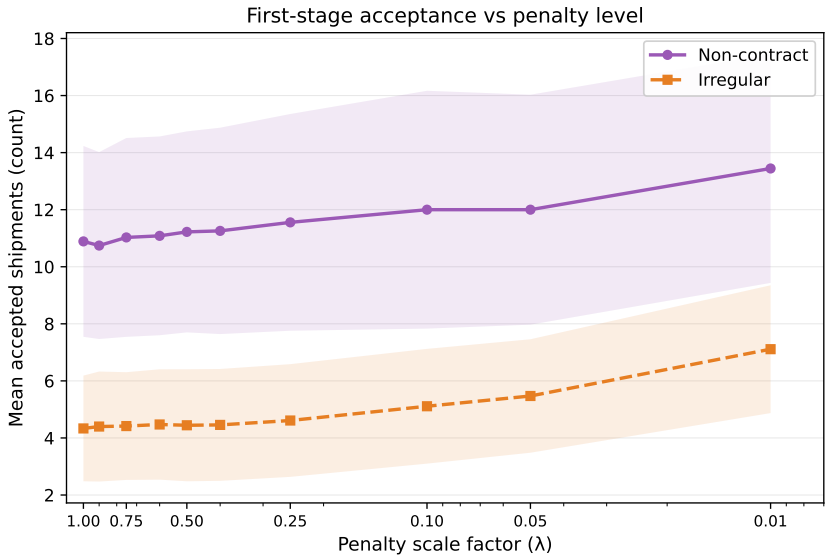


Figure 10: Platform response to penalty reductions. Mean number of accepted non-contract and irregular shipments as a function of λ. Shaded bands show one standard deviation across instances.

7 Conclusions and Perspectives

This paper addressed the tactical planning of M1M systems under simultaneous uncertainty in demand and supply. The IDSP commits to shipment requests, service offers, and shipment itineraries before the season, when the volumes of the requests and the capacities of the services, which it does not control, remain uncertain. To the best of our knowledge, this is the first study to address this problem, extending the work of Taherkhani et al. (2025), which considered demand-side uncertainty only. Our objectives were to formulate a model capturing both sources of uncertainty, to provide a solution approach for instances that cannot be solved directly, and to assess the value of accounting for supply-side uncertainty in tactical planning.

We proposed a two-stage stochastic program, the SSND-UDS, in which, unlike classical stochastic SND models, both service selection and shipment routing are first-stage decisions, reflecting the commitments the IDSP makes to carriers and shippers in advance of the season. We showed that the recourse problem admits an equivalent MKP reformulation, which enables the efficient evaluation of fixed first-stage decisions. For larger instances, we extended the decision-based scenario clustering framework of Hewitt et al. (2022) to joint demand and supply uncertainty, which provides lower and upper bounds that allow the quality of a tactical plan to be assessed when the extensive form cannot be solved within practical time limits. The inverse partition yields tighter bounds than the PAM partition, at a lower computational cost.

The computational results show that stochastic planning increases expected profit over deterministic planning, and that this gain grows with the level of uncertainty. Supply uncertainty contributes approximately 15 times as much to this gain as demand uncertainty. For a platform that relies on capacity it does not control, capacity uncertainty is thus the main source of risk for the tactical plan and should be modeled explicitly. The stochastic plan obtains its advantage through selective commitment, accepting fewer non-contract and irregular shipments and relying more on contract carriers, so that accepting less demand at the tactical level yields a higher expected profit. The results also inform the design of the non-fulfillment penalty, which the IDSP sets before the season. The penalty should be kept above a threshold to maintain a high level of service, set higher for irregular than for non-contract shipments to equalize their service levels, and reviewed at the start of each season, as non-fulfillment increases with the level of uncertainty.

Several perspectives follow from this work. First, the uncertainty considered represents business-as-usual variability around announced volumes and capacities, for which a tactical plan applied throughout the season is appropriate. Major disruptions, such as the prolonged unavailability of a carrier or a terminal, fall outside this representation. Extending the modeling of uncertainty to the resilience of the plan and its recovery after such disruptions is therefore a relevant direction. Second, shipment volumes and

service capacities are sampled independently, whereas Lium et al. (2007, 2009) showed that correlations among demands affect the structure and the robustness of stochastic service network designs. Incorporating correlations among volumes, among capacities, and between the two is a natural extension. Third, the direct solve reaches the time limit on most of the larger instances, and the gap of the cluster bounds is largest under high uncertainty. Decomposition methods, such as the Benders decomposition applied by Taherkhani et al. (2025) or the progressive hedging-based metaheuristics of Crainic et al. (2011), are promising directions for addressing larger instances.

References

I. Bakir, A. Erera, and M.W.F. Savelsbergh. Motor carrier service network design. In T.G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications to Transportation and Logistics*, chapter 14, pages 427–467. Springer, Cham, 2021.

I. C. Bilegan, L. Brotcorne, D. Feillet, and Y. Hayel. Revenue management for rail container transportation. *EURO Journal on Transportation and Logistics*, 4(2):261–283, 2015.

I. C. Bilegan, T. G. Crainic, and Y. Wang. Scheduled service network design with revenue management considerations and an intermodal barge transportation illustration. *European Journal of Operational Research*, 300(1):164–177, 2022.

J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer, New York, 1997.

Bureau of Transportation Statistics. Moving goods in the United States. `https://data.bts.gov/stories/s/Moving-Goods-in-the-United-States/bcyt-rqmu/`. Freight Facts and Figures, "Freight Movement" section. Accessed: October 24, 2026.

M. Chouman and T. G. Crainic. Freight railroad service network design. In T. G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications to Transportation and Logistics*, chapter 13, pages 383–426. Springer, Cham, 2021.

M. Christiansen, K. Fagerholt, B. Nygreen, and D. Ronen. Maritime transportation. In C. Barnhart and G. Laporte, editors, *Transportation*, volume 14 of *Handbooks in Operations Research and Management Science*, pages 189–284. North-Holland, Amsterdam, 2007.

M. Christiansen, E. Hellsten, D. Pisinger, D. Sacramento, and C. Vilhelmsen. Liner shipping network design. *European Journal of Operational Research*, 286(1):1–20, 2020.

J.-F. Cordeau, P. Toth, and D. Vigo. A survey of optimization models for train routing and scheduling. *Transportation Science*, 32(4):380–404, 1998.

T. G. Crainic. Long-haul freight transportation. In R. W. Hall, editor, *Handbook of Transportation Science*, pages 433–491. Kluwer Academic Publishers, Norwell, MA, 1999.

T. G. Crainic. Service network design in freight transportation. *European Journal of Operational Research*, 122(2):272–288, 2000.

T. G. Crainic. Service network design for consolidation-based transportation: Advanced topics. Technical Report CIRRELT-2024-31, CIRRELT, Montréal, Canada, 2024. Forthcoming as Chapter 4 in S. N. Parragh and T. Van Woensel (eds.), Research Handbook on Transport Modeling, Edward Elgar, Cheltenham.

T. G. Crainic and K. H. Kim. Intermodal transportation. In C. Barnhart and G. Laporte, editors, *Transportation*, volume 14 of *Handbooks in Operations Research and Management Science*, chapter 8, pages 467–537. North-Holland, Amsterdam, 2007.

T. G. Crainic and W. Rei. 50 years of Operations Research for the tactical planning of consolidation-based freight transportation. *EURO Journal on Transportation and Logistics*, 14:100157, 2025.

T. G. Crainic, M. Gendreau, and J.-Y. Potvin. Intelligent freight-transportation systems: Assessment and the contribution of operations research. *Transportation Research Part C: Emerging Technologies*, 17(6):541–557, 2009.

T. G. Crainic, X. Fu, M. Gendreau, W. Rei, and S. W. Wallace. Progressive hedging-based metaheuristics for stochastic network design. *Networks*, 58(2):114–124, 2011.

B. Crevier, J.-F. Cordeau, and G. Savard. Integrated operations planning and revenue management for rail freight transportation. *Transportation Research Part B: Methodological*, 46(1):100–119, 2012.

B. Feng and Q. Ye. Operations management of smart logistics: A literature review and future research. *Frontiers of Engineering Management*, 8:344–355, 2021.

W. Guo, B. Atasoy, W. B. van Blokland, and R. R. Negenborn. Global synchromodal transport with dynamic and stochastic shipment matching. *Transportation Research Part E: Logistics and Transportation Review*, 152:102404, 2021.

J. A. Hartigan and M. A. Wong. A K-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1):100–108, 1979.

M. Hewitt, T. G. Crainic, M. Nowak, and W. Rei. Scheduled service network design with resource acquisition and management under uncertainty. *Transportation Research Part B: Methodological*, 128:324–343, 2019.

M. Hewitt, W. Rei, and S. W. Wallace. Stochastic network design. In T. G. Crainic, M. Gendreau, and B. Gendron, editors, *Network Design with Applications to Transportation and Logistics*, chapter 10, pages 283–315. Springer, Cham, 2021.

M. Hewitt, J. Ortmann, and W. Rei. Decision-based scenario clustering for decision-making under uncertainty. *Annals of Operations Research*, 315(2):747–771, 2022.

B. Kam, K. Ahsan, H. Deng, and S. Duan. Digital platforms in supply chain and logistics management: A systematic literature review. *Transportation Research Part E: Logistics and Transportation Review*, 209:104737, 2026.

L. Kaufman and P. Rousseeuw. Partitioning around medoids (program PAM). In *Finding Groups in Data: An Introduction to Cluster Analysis*, chapter 2, pages 68–125. John Wiley & Sons, New York, 1990. doi: 10.1002/9780470316801.ch2.

J. Keutchayan, J. Ortmann, and W. Rei. Problem-driven scenario clustering in stochastic optimization. *arXiv preprint arXiv:2106.11717*, 2021.

A. J. Kleywegt, A. Shapiro, and T. Homem-de-Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2): 479–502, 2002.

W. Klibi, A. Martel, and A. Guitouni. The design of robust value-creating supply chain networks: a critical review. *European Journal of Operational Research*, 203(2):283–293, 2010.

A.-G. Lium, T. G. Crainic, and S. W. Wallace. Correlations in stochastic programming: A case from stochastic service network design. *Asia-Pacific Journal of Operational Research*, 24(2):161–179, 2007.

A.-G. Lium, T. G. Crainic, and S. W. Wallace. A study of demand stochasticity in service network design. *Transportation Science*, 43(2):144–157, 2009.

M. SteadieSeifi, N. P. Dellaert, W. Nuijten, T. Van Woensel, and R. Raoufi. Multimodal freight transportation planning: A literature review. *European Journal of Operational Research*, 233(1):1–15, 2014.

G. Taherkhani, I. C. Bilegan, T. G. Crainic, M. Gendreau, and W. Rei. Tactical capacity planning in an integrated multi-stakeholder freight transportation system. *Omega*, 110: 102628, 2022.

G. Taherkhani, M. Hosseini, and A. Hassanzadeh. Exact solution method for multi-stakeholder freight transportation systems under uncertainty. *Transportation Research Part B: Methodological*, 200:103288, 2025. doi: 10.1016/j.trb.2025.103288.

B. van Riessen, R. R. Negenborn, and R. Dekker. The cargo fare class mix problem for an intermodal corridor: revenue management in synchromodal container transportation. *Flexible Services and Manufacturing Journal*, 29(3):634–658, 2017.

B. Van Riessen, J. Mulder, R. R. Negenborn, and R. Dekker. Revenue management with two fare classes in synchromodal container transportation. *Flexible Services and Manufacturing Journal*, 33(3):623–662, 2021.

Y. Wang, I. C. Bilegan, T. G. Crainic, and A. Artiba. Performance indicators for planning intermodal barge transportation systems. *Transportation Research Procedia*, 3:621–630, 2014.

Y. Wang, I. C. Bilegan, T. G. Crainic, and A. Artiba. A revenue management approach for network capacity allocation of an intermodal barge transportation system. In *Computational Logistics. ICCL 2016*, volume 9855 of *Lecture Notes in Computer Science*, pages 243–257, Cham, 2016. Springer.

Y. Zhao, Q. Xue, Z. Cao, and X. Zhang. A two-stage chance constrained approach with application to stochastic intermodal service network design problems. *Journal of Advanced Transportation*, 2018:6051029, 2018. doi: 10.1155/2018/6051029.

A Experimental results with Scenario Clustering

A.1 Selection of the Number of Clusters by Silhouette Score

This section examines whether the silhouette score of the PAM partition can guide the choice of κ. The analysis is restricted to the PAM partition, since the silhouette score measures the separation between homogeneous clusters, whereas the clusters of the inverse partition are heterogeneous by construction. Gaps are computed as $100 \times (OPT_{\text{Direct}} - \Upsilon)/OPT_{\text{Direct}}$, where Υ denotes the bound considered.

Figure 11 relates the silhouette score to bound tightness across $\kappa \in \{10, 20, 30, 40\}$. For both lower bounds, the fitted trends are mildly negative, indicating that better-separated partitions are associated with tighter lower bounds, although the dispersion around the fitted lines is substantial. For Cluster-UB, the fitted trends are essentially flat, which indicates that the looseness of the upper bound depends on the number of clusters rather than on their separation.

We then assess whether this association yields a useful selection rule. For each instance, κ is set to the value maximizing the average silhouette score of the PAM partition over $\{10, 20, 30, 40\}$, referred to as best κ. Over the 61 instances for which the direct solve proves optimality, the rule selects $\kappa = 10$, 20, 30, and 40 for 15, 12, 20, and 14 instances, respectively. Table 14 reports the resulting bounds.

Table 14: Summary statistics (in %) for the bounds on the PAM partition when κ is selected per instance as the value maximizing the silhouette score over $\{10, 20, 30, 40\}$ (best κ), on the subset of instances where the direct solve proves optimality ($LB_{\text{Direct}} = UB_{\text{Direct}}$, 61 instances). Worst Gap is the gap of largest absolute value.

Measure	Mean Gap	Median Gap	Std Gap	Worst Gap
Medoid-LB (best κ)	2.29	1.04	3.60	21.38
Cluster-LB (best κ)	2.46	1.03	3.23	15.38
Cluster-UB (best κ)	−6.08	−4.78	4.56	−21.97

The mean gap of Medoid-LB under best κ is 2.29%, tighter than at $\kappa = 10$ (4.04%) but looser than at $\kappa = 30$ (1.46%) and $\kappa = 40$ (1.16%), as reported in Table 7. For the cluster bounds, the rule is detrimental. The mean gap of Cluster-LB is 2.46%, against 2.03% at $\kappa = 10$ on the same partition, and the mean gap of Cluster-UB is −6.08%, against −4.55% at $\kappa = 10$. Neither best-κ cluster bound approaches the corresponding bound on the inverse partition at $\kappa = 10$ (0.93% and −2.19%). The weak association observed in Figure 11 therefore does not translate into a reliable criterion for selecting κ when the objective is bound certification, and the number of clusters is fixed at $\kappa = 10$ in the main experiments.

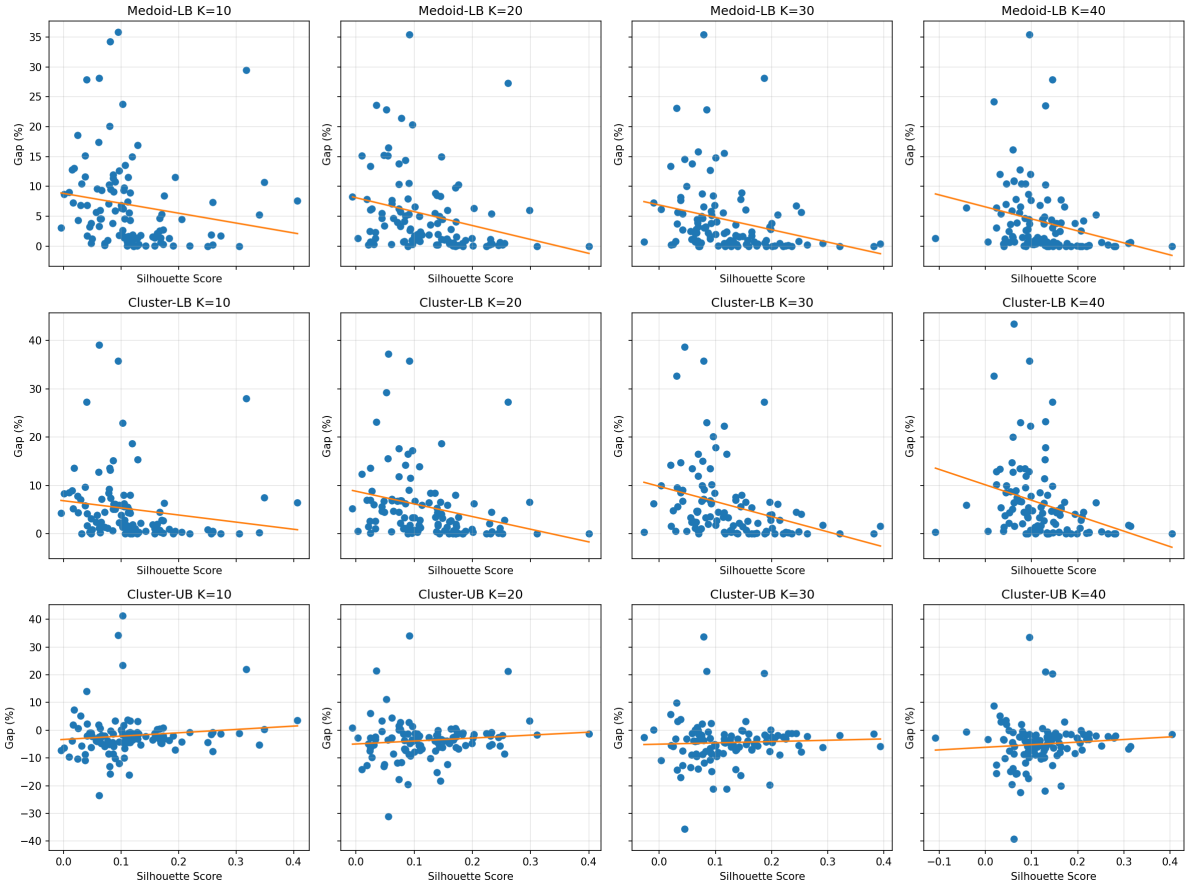


Figure 11: Scatterplots of silhouette score versus percentage gap for Medoid-LB (top row), Cluster-LB (middle), and Cluster-UB (bottom), across $\kappa \in \{10, 20, 30, 40\}$. Orange lines are least-squares fits. Plots are shown for instances with complete silhouette data for all κ values (105 instances after filtering).

A.2 Impact of Instance Size on Run-Time (Direct Solve and Scenario Distance Calculation)

All instances considered here have a fixed number of scenarios ($|\mathcal{S}| = 100$), so we stratify by (Terminals, Shipments, Services). Table 15 reports (i) the direct solve time and resulting optimality gap under a fixed wall-clock time limit of 24 hours, and (ii) the scenario distance–matrix computation time used by clustering. The table aggregates over all instances with complete timing data.

Table 15 confirms that direct solving becomes rapidly prohibitive as the number of terminals, shipments, and services increase: many of the larger groups hit the time limit and exhibit both higher mean gaps and fewer instances with zero gap. In contrast, the distance-matrix computation remains consistently feasible across all groups, though its

Table 15: Effect of instance size on direct (Gurobi) solve and scenario–distance computation for all instances with $|\mathcal{S}| = 100$.

Inst.			Direct Solve					Dist. Calc.	
			Time (s)			(%)		Time (s)	
Term.	Ship.	Serv.	Mean	Std	Med.	Mean Gap	Std Gap	Mean	Std
10	50	2	530.0	258.5	434.8	0.00	0.00	1058.5	102.0
10	50	4	6321.2	8953.2	643.3	0.00	0.00	1593.3	284.3
10	80	2	46916.7	31042.7	52920.1	0.14	0.43	1519.7	129.7
10	80	4	31110.6	9322.0	32035.4	0.00	0.00	2177.6	181.0
10	100	2	63973.3	28573.1	86224.9	5.55	13.39	1858.3	40.2
10	100	4	86226.1	2.6	86225.2	3.07	3.32	2296.6	200.6
20	50	2	32708.6	29590.2	19716.1	0.29	0.86	4135.9	221.2
20	50	4	73249.0	24950.8	86254.8	4.31	4.11	6544.3	359.5
20	80	2	86276.9	12.7	86271.8	10.64	16.42	5639.9	652.8
20	80	4	86291.9	5.8	86290.2	12.62	7.44	8470.1	249.1
20	100	2	86291.9	7.0	86289.3	12.23	16.00	6252.9	126.4
20	100	4	86309.9	10.1	86313.6	24.82	15.25	9644.2	733.6
30	50	2	19667.1	17691.1	17157.0	0.00	0.00	9648.8	785.2
30	50	4	64316.6	33684.8	86326.9	2.22	2.36	16453.0	1932.4
30	80	2	81484.0	10056.2	86334.7	14.56	21.91	11310.4	933.3
30	80	4	86384.3	10.7	86381.5	3.18	0.90	19179.1	1384.1
30	100	2	86394.9	3.6	86395.0	21.34	6.32	13498.1	701.0

cost increases materially with instance size. These results motivate the use of clustering-based surrogates for large instances while retaining the direct model as the benchmark when it can be solved to high quality.

B Detail Report of Value of Stochastic Solution Experiments

Table 16: Instance-level VSS results by network size and uncertainty level (CV). Rep. indicates the replicate index.

Periods	Terminals	Shipments	Serv./pair	CV	Rep.	RP	EEV	VSS (%)
7	10	50	2	0.15	0	4195.27	3642.49	13.18
7	10	50	2	0.15	1	2830.34	2741.02	3.16
7	10	50	2	0.15	2	4121.34	3930.93	4.62
7	10	50	4	0.15	0	2688.98	2677.98	0.41
7	10	50	4	0.15	1	4893.97	4893.97	0.00
7	10	50	4	0.15	2	2619.50	2226.26	15.01
7	20	50	2	0.15	0	2319.83	2319.83	0.00
7	20	50	2	0.15	1	2045.15	2045.15	0.00
7	20	50	2	0.15	2	2889.36	2889.36	0.00
7	20	50	4	0.15	0	3998.88	3998.88	0.00
7	20	50	4	0.15	2	2372.22	2372.22	0.00
7	30	50	2	0.15	0	2523.32	2523.32	0.00
7	30	50	2	0.15	1	4745.83	4745.83	0.00
7	30	50	4	0.15	1	2725.17	2717.12	0.30
7	10	50	2	0.25	0	4137.93	3613.82	12.67
7	10	50	2	0.25	1	2735.93	2490.96	8.95
7	10	50	2	0.25	2	4007.64	3599.65	10.18
7	10	50	4	0.25	0	2656.78	2325.90	12.45
7	10	50	4	0.25	1	4826.26	4826.26	0.00
7	10	50	4	0.25	2	2579.28	1888.85	26.77
7	20	50	2	0.25	0	2316.50	2316.50	0.00
7	20	50	2	0.25	1	2100.27	2053.27	2.24
7	20	50	2	0.25	2	2792.95	2783.11	0.35
7	20	50	4	0.25	0	4009.43	4009.43	0.00
7	30	50	2	0.25	0	2501.33	2497.11	0.17
7	30	50	2	0.25	1	4717.02	4717.02	0.00
7	30	50	2	0.25	2	3020.62	3016.64	0.13
7	30	50	4	0.25	1	2724.30	2693.25	1.14
7	10	50	2	0.35	0	4039.31	3366.03	16.67
7	10	50	2	0.35	1	2598.93	2291.84	11.82
7	10	50	2	0.35	2	3619.46	3097.44	14.42
7	10	50	4	0.35	0	2517.47	1952.82	22.43
7	10	50	4	0.35	1	4756.04	4726.20	0.63
7	10	50	4	0.35	2	2164.70	1020.58	52.85
7	20	50	2	0.35	0	2227.93	2227.93	0.00
7	20	50	2	0.35	1	1934.93	1781.90	7.91
7	30	50	2	0.35	0	2452.68	2417.33	1.44
7	30	50	2	0.35	1	4675.06	4654.18	0.45
7	30	50	2	0.35	2	2875.84	2708.71	5.81
7	30	50	4	0.35	1	2622.63	2317.45	11.64
7	10	100	2	0.15	0	7576.59	6264.01	17.32
7	10	100	2	0.15	2	7995.41	7272.90	9.04
7	10	80	2	0.15	0	5775.29	4928.95	14.65
7	10	80	2	0.15	1	4999.19	4963.60	0.71
7	10	80	2	0.15	2	7776.87	7396.64	4.89
7	10	80	4	0.15	0	5952.39	5873.19	1.33
7	10	80	4	0.15	1	4679.95	4659.10	0.45
7	30	80	2	0.15	0	5750.17	5750.17	0.00
7	10	100	2	0.25	2	7722.23	6434.59	16.67
7	10	80	2	0.25	0	5566.21	4290.77	22.91
7	10	80	2	0.25	1	4760.21	4507.55	5.31
7	10	80	2	0.25	2	7480.17	6790.99	9.21
7	10	80	4	0.25	0	5850.23	5379.55	8.05
7	10	80	4	0.25	1	4614.43	4476.83	2.98
7	10	80	4	0.25	2	6540.77	5290.85	19.11
7	10	100	2	0.35	2	7163.29	5508.90	23.10
7	10	80	2	0.35	0	5239.35	3312.08	36.78
7	10	80	2	0.35	2	6968.07	5481.48	21.33
7	10	80	4	0.35	0	5615.39	4918.00	12.42
7	10	80	4	0.35	1	4201.44	3443.40	18.04
7	30	80	2	0.35	0	5535.74	5329.96	3.72