CIRRELT

CIRRELT-2025-16

Service Network Design with Uncertainty on Water Levels for Intermodal River Transport

Bita Payami
Ioana C. Bilegan
Teodor Gabriel Crainic
Walter Rei

May 2025

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

Service Network Design with Uncertainty on Water Levels for Intermodal River Transport

Bita Payami[1], Ioana C. Bilegan[2], Teodor Gabriel Crainic[1,*], Walter Rei[1]

1. Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), School of Management, Université du Québec à Montréal
2. Univ. Polytechnique Hauts-de-France, LAMIH, CNRS, UMR 8201, F-59313 Valenciennes, France

Abstract. Barge transport presents a sustainable alternative to road transport, offering the potential to alleviate congestion and reduce costs. However, increasingly frequent and severe drought seasons, attributed to climate change, challenge the resilience of this mode of transport. Decreased water levels restrict navigation, impact vessel size, and reduce vessel capacity. This study introduces a novel and integrated modeling framework for tactical planning in consolidation-based barge freight transportation. By examining the relationship between water levels and vessel load capacity, with a specific focus on vessel dimensions, the framework aims to evaluate possible impacts on the efficiency and profitability of inland waterway transport. We also introduce two stochastic Water-Level-Constrained Scheduled Service Network Design with Resource and Revenue Management models to address uncertainties in resource capacities caused by water level variations, each addressing a distinct tactical plan adjustment strategy. These models aim to establish a tactical plan, given predicted water levels, that maximizes the expected carrier's revenue while accounting for future adjustments to the plan when information is revealed and predictions are reliably updated, to fulfill the demands of shippers and optimize the utilization of the carrier's resources. Through extensive experimentation using commercial software and a novel decision-based scenario clustering algorithm, we assess the quality of the solutions obtained with the different model variants and analyze the impact of the water level on the results.

Keywords: Consolidation-based barge transportation; two-stage stochastic programming; water level fluctuations; tactical planning; scheduled service network design; scenario reduction; clustering.

Acknowledgments: While working on the project, the first author was Ph.D. candidate at the School of Management, Université du Québec à Montréal (UQAM), Canada, and member of CIRRELT. The third author held the UQAM Chair in Intelligent Logistics and Transportation Systems Planning, and was Adjunct Professor in the Department of Computer Science and Operations Research at the Université de Montréal, while the fourth author held the Canada Research Chair in Stochastic Optimization of Transport and Logistics Systems. We gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Collaborative Research and Development, and Discovery grant programs. We also gratefully acknowledge the support of Fonds de recherche du Québec through their infrastructure grants. Finally, during the preparation of this work the authors used ChatGPT-4, under its most strict privacy enforcing settings, for text edit purposes exclusively. After using this tool, the authors reviewed and further edited the text as needed and take full responsibility for the content of the publication.

---

* Corresponding author:teodorgabriel.crainic@cirrelt.net

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2025

1 Introduction

Barge transportation, a key component of intermodal freight transportation, is known for its cost-effectiveness and environmental sustainability and plays a crucial role in facilitating the exchange of freight between maritime ports and their hinterlands, as well as among river ports. Carriers in this sector often use consolidation strategies to combine smaller shipments, enhancing operational efficiency and cost-effectiveness. However, this requires precise coordination of shipping schedules, freight specifications, and service requirements. To deliver shipments at low cost and with high punctuality, carriers need effective planning of transportation activities—detailing the service network, resource utilization, and shipment handling and transport operations—in a way that ensures efficiency, profitability, and effective consolidation, while maintaining the desired level of service quality that is crucial for shipper satisfaction. This problem, addressed at the tactical level, is supported by the Service Network Design (SND) methodology. This tactical planning problem is typically addressed using the Scheduled Service Network Design (SSND) model, which defines an optimized transportation plan that specifies which services and schedules to operate, and how shipments are routed across the network for a given schedule length (e.g., a week), executed repeatedly over a medium-term horizon (e.g., a season), with the objective of maximizing profit.

Most service network design cases, including SSND models, often overlook environmental factors that can directly impact the capacity of transportation resources, which are used for operating scheduled services efficiently across land-, air-, and water-based modes. Beyond disregarding infrastructure conditions, these models are predominantly deterministic, operating under the assumption that there are no significant variations in the system's state—whether on the supply or demand sides—throughout the planning horizon (Crainic and Hewitt, 2021; Crainic and Rei, 2025). SSND models generally presume that the capacity offered by each resource supporting the service over the designed network remains known and constant throughout the planning horizon. This is often not the case, however. Hence, our research centers on the uncertainty in infrastructure and resource capacity, particularly relevant for inland waterway transportation, where water-level fluctuations can significantly impact vessel capacity. Lower water levels can decrease vessel capacity due to increased grounding risks, while higher water levels might allow for greater freight capacity but could introduce navigational challenges under bridges and through certain canal sections (Prandtstetter et al., 2023). These fluctuations result from transient weather conditions, such as rainfall and short dry spells, and are characterized as randomness—foreseeable variations that can be described through random variables with known probability distributions (Christodoulou et al., 2020; Zheng and Kim, 2017). Expert-generated forecasts provide probability distributions for water levels at critical river segments and ports, specifying the expected variation in a vessel's loading capacity based on its type. Such predictions are made for the next season at a fairly high aggregation level and are later repeatedly updated during operations for short (e.g., the schedule length) horizons. When the updated predicted water levels are lower than initially anticipated, carriers face reduced vessel capacities, forcing them to either reject part of the demand to be transported and pay any associated penalties or adapt their transportation plans through costly re-optimization. These dynamic adjustments, which deterministic SSND models are unequipped to handle, underscore the need for a stochastic service network design framework that integrates infrastructure uncertainty and enables flexible, informed decision-making under environmental variability. While

there is a substantial body of literature on SSND models for consolidation-based transportation, most studies focus on demand-side uncertainty, with only a few addressing supply-side uncertainty and even fewer considering uncertainties in infrastructure and resource capacity. To the best of our knowledge, no existing work has jointly tackled both the design of an efficient service network and the impact of water level uncertainty on vessel-supported service capacity.

We aim to fill this gap by introducing the *Water-Level-Constrained Scheduled Service Network Design with Resource and Revenue Management*(WL-SSND-RRM) problem, which addresses the challenges of integrating water-level variability into tactical planning for barge transportation systems. We propose two models, formulated within a two-stage stochastic programming framework, each defined by a specific strategy to handle the impact of randomly changing water-level conditions. Both models aim to establish a tactical operations plan, given predicted water levels, that maximizes the expected carrier's revenue while accounting for future adjustments to the plan as information is revealed and forecasts are updated—thereby fulfilling shipper demand and optimizing resource utilization.

In the proposed stochastic models, water level uncertainty is represented through a scenario set that captures random variations in service capacities. To reflect this variability accurately, the scenario set must be sufficiently comprehensive, which often requires a large number of scenarios and, consequently, increases the amount of contextual information that must be integrated into the optimization process. However, incorporating all scenarios into the stochastic WL-SSND-RRM models can significantly increase computational complexity, rendering them intractable in practice. To address this challenge, we propose a two-step solution approach. In Step 1, we interpret the scenarios in terms of the decisions they induce and form clusters of scenarios that lead to similar decisions, reducing redundancy in the scenario set and simplifying it while minimizing information loss. Additionally, we form clusters of scenarios that lead to dissimilar decisions, ensuring that each cluster captures diverse, decision-relevant information. In Step 2, we leverage the scenario clusters obtained in Step 1 to apply a series of bounding techniques, providing both a set of alternative feasible solutions, whose objective function values define alternative lower bounds for the stochastic models under consideration, and an upper bound value, which serves as a bounding certificate for evaluating the optimality gaps of the obtained feasible solutions.

The contributions of this paper are the following: (i) We introduce the WL-SSND-RRM problem, which explicitly accounts for uncertainties in resource capacities caused by random water-level variations—a critical factor that directly impacts vessel capacity and service feasibility but has not been addressed in the literature; (ii) This work propose two stochastic models, each defined by a specific approach to handling the impact of randomly changing water level conditions. The first model introduces penalties for unmet demand, maintaining a fixed tactical plan even when water levels deviate from expectations. The second model allows for demand itinerary adjustments at a cost, enabling shipment flows to adapt dynamically to observed fluctuations in water levels; (iii) We apply a decision-based scenario clustering analysis approach—tailored to our problem setting—that enables the computation of a series of alternative bounding techniques aimed at finding feasible solutions and computing the optimality gap; (iv) We conduct extensive computational experiments to evaluate the performance of our stochastic models compared to the deterministic model. Also, we analyze how different strategies to handling the impact of randomly changing water levels influence the structural

characteristics of tactical plans. Additionally, we assess the impact of water level fluctuations on efficiency and profitability.

The paper is organized as follows: Section 2 presents the barge transportation system, including the physical network, supply side, and demand side. Section 3 provides a comprehensive review of the relevant literature, encompassing discussions on both deterministic and stochastic service network design, and examines various solution methods proposed in this field. Section 4 outlines the problem setting and introduces the proposed stochastic WL-SSND-RRM formulations for the barge transportation system. Section 5 explains the solution approach adopted in this research. Section 6 reports the experimental results. Finally, Section 7 summarizes the main findings and outlines future research directions.

2 Barge transportation system description

Freight transportation via inland waterways involves the movement of shipments from their point of origin to their destination using vessels that operate along rivers, canals, and related infrastructure (Bilegan et al., 2022). A detailed description of the physical infrastructure involved in barge transportation, including waterways and terminals, will be provided in Section 2.1. The key stakeholders in this system are carriers and shippers, representing the supply and demand sides, respectively, which will be further explored in Sections 2.2 and 2.3.

2.1 Physical network

Similar to land- and air-based transportation modes, which require a certain set of infrastructure components, water-based transportation has its own physical network, which consists of waterways and ports. Waterways are corridors characterized by their length, width, and depth, connecting ports and allowing vessels to navigate between them to transport freight from one terminal to another. Depending on their geographic location, waterways are categorized as deep sea, short sea, coastal, or inland. In contrast to the deep sea, the coastal and inland waterways studied here are shallower, narrower, and shorter. Due to these physical characteristics, smaller vessels, such as barges, are used to navigate these waterways. These attributes limit not only the size of vessels but also the number of vessels that can travel simultaneously in both directions on a given section of the waterway within a specific time frame. While the width and length of rivers remain relatively constant over time, the depth—defined as the distance between the water surface and the highest point of the riverbed—fluctuates. Water depth is influenced by various factors over short, medium, and long-term periods, such as extreme weather events, seasonal changes, and annual warming and cooling cycles, making navigation more challenging at times.

Ports, the other key component of water-based transportation, are facilities located along coasts or rivers where vessels dock to load and unload freight. Ports are composed of one or more quays and terminals. A quay is the structure along the edge of the water where vessels are moored, and its length is often considered to match that of the terminal, especially in simplified terminal layouts. A terminal is a specialized facility within a port designed for handling freight. Terminals are equipped with cranes, warehouses, storage areas, and other

infrastructure necessary for loading and unloading vessels and transferring freight to and from trucks, trains, or other modes of transportation. The main physical constraints of each terminal are the water depth at the berth, the number of load-handling machines, and the available storage space. The berthing capacity—the number of vessels that can dock simultaneously—is limited by the terminal's length. The water depth limits the maximum submerged vertical dimension of a loaded vessel, defining the safe docking depth at the terminal, while storage capacity limits the volume of freight that can be stored at the terminal for a given period. Therefore, each terminal's capacity is measured based on several dimensions, such as the volume of freight that can be loaded or unloaded per time unit (i.e., handling capacity), the number of vessels that can be docked at the terminal in a given length unit (i.e., berthing capacity), and the volume of freight that can be stored at the terminal per time unit (i.e., storage capacity). Carriers using the port have to pay the costs associated with each terminal. Load handling, load storage, and vessel holding costs are all included in the costs associated with each terminal. The first two costs—load handling cost and load storage cost—vary depending on the type of container used, while the vessel holding cost for each terminal depends on the type of the vessel.

2.2 The barge transportation system's supply side

Carriers in the barge transportation system offer scheduled services to transport freight between terminals. These services are designed to meet a variety of shipping needs by offering flexible routes and schedules to ensure shipments are delivered from origin to destination within a specified time window. Each potential scheduled service follows a specific waterway route within the network, which may be either a single-leg service (traveling directly from origin to destination without stops) or a multi-leg service (stopping at multiple terminals to load and unload freight). Each service is defined by its origin, destination, intermediate terminals (if any), and scheduled arrival and departure times at each terminal. The total duration of any service, whether single-leg or multi-leg, is determined by the cumulative travel times for each leg and any waiting times at terminals(Crainic and Hewitt, 2021). Carriers offer a range of fare classes, which are price categories designed to reflect different levels of service quality that may vary in terms of delivery time and flexibility. These fare classes allow shippers to choose between different service quality options. For delivery time, shippers can opt for standard delivery, which involves longer transit times, or express delivery, which offers faster transit times. Overall, fare classes enable carriers to provide tailored quality of service levels that meet the varying needs of shippers.

Services are performed by vessels of various types in the barge transportation system. Each vessel type has specific technical specifications: length, typically measured in meters or feet; speed, measured in knots; and critical depth characteristics, such as maximum and minimum draught (See Figure1). The maximum draught refers to the deepest point a fully loaded vessel can safely reach below the waterline, determining the maximum load it can carry without compromising safety. Conversely, the minimum draught represents the shallowest depth the vessel reaches when empty (Prandtstetter et al., 2023). These last two predefined specifications establish the safe loading range, or loading capacity, of each vessel, which can be measured in terms of weight (tons), volume (TEUs), or a combination of both. When a vessel is partially loaded (i.e., carrying less than its maximum capacity but not empty), its submerged depth—referred to as the actual draught—naturally falls between the minimum and maximum draughts, as

illustrated in Figure 1. The actual draught can be described as an incremental function, where each added unit of load results in a proportional increase in draught (van Dorsser et al., 2020). In this context, the load-draught coefficient signifies how sensitive the draught is to changes in load, capturing this relationship whether it follows a simple linear progression or a more complex pattern.

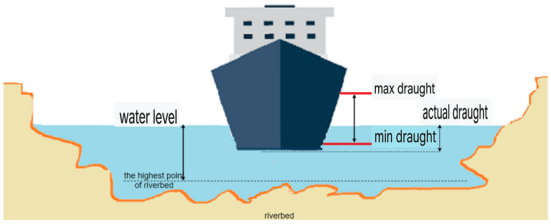


Figure 1: Schematic view of the most relevant measurements related to vessel's draught and water level

2.3 The barge transportation system's demand side

Shippers generate the demand for transportation services in the barge transportation system. These shippers include manufacturers, distributors, wholesalers, retailers, and other businesses that require the movement of shipments from one port to another. The shipments may consist of finished products, intermediate goods, raw materials, or other types of freight, which can be packaged in various forms such as boxes, bags, drums, bales, and rolls. In the intermodal barge transportation system, loads transported by vessels are typically placed into containers, offering several advantages, including improved load safety, reduced handling costs, and standardization (Bektaş and Crainic, 2008). Standard containers come in two lengths—20 feet and 40 feet. The Twenty-foot Equivalent Unit (TEU) is the standard measure of container volume, where a 20-foot container equals 1 TEU, and a 40-foot container equals 2 TEUs.

Shippers on the demand side of the barge transportation system can be categorized based on their business relationship with carriers as regular or spot shippers; Regular shippers: have long-term contracts with carriers, ensuring consistent and predictable transportation demands and benefiting from guaranteed capacity. Spot shippers: on the other hand, request transportation services only when needed, without engaging in long-term contractual commitments (Bilegan et al., 2022). Regardless of whether a shipper is regular or spot, each shipper's demand has several defining characteristics: the origin and destination of the shipment, specifying the pickup and delivery ports; the demand size, measured by volume (TEUs) and weight (total tonnage); and the type of container, such as open-top, open-side, or refrigerated. Additionally, shippers choose fare options based on their delivery needs, opting for either standard delivery or express delivery.

3 Literature Review

This section reviews the relevant literature on planning problems in intermodal barge transportation systems. The discussion emphasizes three interconnected aspects: planning studies related to service network design in barge transportation, the influence of infrastructure and environmental limitations on tactical decisions, and the development of optimization-based models and solution techniques proposed to support intermodal barge transportation at the tactical planning level.

Service network design problems have been extensively studied in the literature to address tactical planning issues for consolidation-based carriers (Crainic and Hewitt, 2021; Crainic and Rei, 2025). The reader may refer to reviews on this field for long-haul transportation by (Crainic, 2003), for rail by (Cordeau et al., 1998), for maritime transportation (Christiansen et al., 2007), for motor carriers (Bakir et al., 2021) and for intermodal transportation by (Crainic and Kim, 2007; Agamez-Arias and Moyano-Fuentes, 2017; SteadieSeifi et al., 2014). Despite the breadth of research on service network design, studies specifically targeting tactical planning in the context of barge transportation are relatively scarce. In intermodal barge transportation, tactical planning has been studied with a focus on service route selection, service frequency and capacity definitions, and in some cases, the repositioning of empty containers (Caris et al., 2012; Braekers et al., 2013; Riessen et al., 2015; Ypsilantis and Zuidwijk, 2019). Nevertheless, most of these studies concentrate on barge routing and repositioning without incorporating broader operational challenges, particularly those related to water depth variability and bridge height limitations—key factors that directly affect the feasibility of barge operations. Only a few studies explicitly consider these infrastructure-related constraints. A notable contribution is by (Zhang et al., 2020), who propose the use of foldable containers to address height and depth limitations in river–sea intermodal transport. Their mixed-integer linear programming model optimizes empty container repositioning under these structural constraints in a deterministic setting. (Müllerklein and Fontaine, 2025) also work on integrating disruption risks into transportation planning by proposing a two-stage stochastic model that optimizes the selection and combination of resilience strategies—such as rerouting, inventory positioning, and sourcing flexibility—under transport uncertainty. Although their work focuses on general supply chain networks, the approach highlights the importance of incorporating tactical and operational responses to infrastructure-related disruptions. This perspective is relevant for barge transportation systems, where fluctuating water levels can significantly impact route feasibility and vessel capacity, yet remain largely overlooked in the existing tactical planning literature.

A significant share of the literature on tactical planning for consolidation-based carriers leverages time-dependent service network design models, commonly referred to as Scheduled Service Network Design (SSND) models. These models incorporate the temporal dimensions of both demand and service through time-space networks, which are constructed by extending the network along the dimension of time for the duration of the schedule length. SSND models have been extensively applied in deterministic frameworks across transportation modes. Comprehensive overviews of SSND models are presented in (Crainic and Hewitt, 2021) and (Crainic and Rei, 2025), including model structures, applications, and solution strategies. SSND has recently been adapted to the context of intermodal barge transportation. (Bilegan et al., 2022) developed a model tailored for inland waterway networks, enhancing the classical SSND framework by integrating resource and revenue management—resulting in the Scheduled Service

Network Design with Resource and Revenue Management (SSND-RRM). This model jointly optimizes service selection, resource utilization, and demand routing to maximize carrier profitability while selecting high-value spot shipments and accommodating varied service quality requirements. Building on these contributions, our work develops a tactical planning framework that explicitly incorporates the stochastic nature of vessel-supported service capacity, which is affected by unavoidable water level fluctuations due to environmental and weather-related variability. We formulate two-stage stochastic programming models to capture this uncertainty—an aspect rarely addressed in SSND literature despite its operational significance. While demand uncertainty has been the primary focus of stochastic SSND models (Bai et al., 2014; Hewitt et al., 2019; Hoff et al., 2010; Lium et al., 2009; Wang et al., 2019; Wang and Qi, 2020), other sources of uncertainty remain underexplored. Most models are typically structured as the two-stage stochastic programming framework, with service selection in the first stage under probabilistic demand, and recourse actions such as rerouting or outsourcing in the second stage once actual demand is revealed (Crainic et al., 2011, 2014). Some contributions, such as (Hewitt et al., 2019), further extend this framework by including resource acquisition and allocation decisions in the first stage. In comparison, travel time uncertainty has received limited attention. (Lanza et al., 2021) proposed using penalties for lateness as a recourse strategy, while (Lanza et al., 2024) proposed a more adaptive approach involving departure rescheduling, flow rerouting, and outsourcing. In contrast, capacity uncertainty remains largely overlooked in the SSND literature. To date, (Sun et al., 2018) is the only study that explicitly models this aspect, using fuzzy chance constraints within a Mixed Integer Linear Programming (MILP) framework to capture rail capacity limitations in intermodal routing, ensuring feasibility with a certain probability. Notably, the model does not incorporate recourse decisions to address capacity shortfalls, leaving a significant gap in adaptive planning under resource capacity uncertainty.

As mentioned, two-stage stochastic programming is the most commonly used formulation to model uncertainty in SSND problems. In these models, uncertainty is captured through a set of scenarios representing random variations, which must be large enough to provide an accurate and detailed representation of uncertainty. However, a larger scenario set significantly increases computational complexity in the optimization process. To address this challenge, various solution methods have been explored, including metaheuristics (Hoff et al., 2010), progressive-hedging-based meta-heuristic (Lanza et al., 2021), matheuristics that utilize column generation (Hewitt et al., 2019), and partial Benders decomposition methodology (Crainic et al., 2021). The complexity of these stochastic combinatorial optimization problems has led to the adoption of scenario clustering, where meta-heuristics first cluster scenarios and then apply progressive hedging to solve the subproblems within these clusters, demonstrating computational success in stochastic network design (Crainic et al., 2014; Jiang et al., 2021). Recently, (Hewitt et al., 2022) proposed a solution approach consisting of two steps: first, scenarios are grouped based on their similarity or dissimilarity in the solutions they induce; second, the scenario cluster information is leveraged to compute efficient bounds for stochastic two-stage service network design models.

4 Problem setting and modeling approach

This section presents two stochastic formulations of the Water-Level-Constrained Scheduled Service Network Design with Resource and Revenue Management (WL-SSND-RRM) problem

developed for the tactical planning of barge transportation, where the plan is established for a schedule length and repeated over the planning horizon, referred to as a season. The WL-SSND-RRM problem involves designing a scheduled service network, managing resources to support services, optimizing profit through revenue management via capacity management, and routing demand under fluctuating water levels, ensuring demand fulfillment while maintaining profitability and operational efficiency. First, we introduce the physical system, demand, potential service network, and economic elements that apply to both stochastic variants of the WL-SSND-RRM problem in Section 4.1. This is followed by a detailed explanation of the specific aspects and mathematical formulations for each variant in Section 4.2.

4.1 Problem Setting and Notation

Physical System. The physical infrastructure system is represented by a physical network, $\mathcal{G}^{\text{ph}} = (\mathcal{N}^{\text{ph}}, \mathcal{A}^{\text{ph}})$, where $\mathcal{N}^{\text{ph}}$ denotes the set of geographical ports and $\mathcal{A}^{\text{ph}}$ represents the set of physical links. Each port $\eta \in \mathcal{N}^{\text{ph}}$ has a specific berthing capacity, measured in length units and denoted by Q_η. Each physical link $a \in \mathcal{A}^{\text{ph}}$ is characterized by a specific water level at each time period, denoted by p_a, which can be represented either by a point estimate in the deterministic case or by a probability distribution in the stochastic case. The point estimate is typically derived statistically from historical data, often using the mode or average to represent the most likely water level on each physical link over the planning horizon. In contrast, water levels in the stochastic case are modeled using probability distributions to capture variability and uncertainty. These distributions reflect both the likelihood of water levels falling within a typical range and the less probable extreme values, providing a comprehensive representation of water level fluctuations over time.

Demand. We define the set of shipping demands $\mathcal{K} = \mathcal{K}^{\text{R}} \cup \mathcal{K}^{\text{P}} \cup \mathcal{K}^{\text{F}}$, where $\mathcal{K}^{\text{R}}$ (regular shippers) includes demands that must be fully served, $\mathcal{K}^{\text{P}}$ (partial-spot shippers) includes demands that may be partially met, and $\mathcal{K}^{\text{F}}$ (full-spot shippers) includes demands that can either be fully met or not met at all. Each demand $k \in \mathcal{K}$ is characterized by its volume $d(k)$ in TEUs. The conversion factor from TEUs to tonnage, denoted by ω, relates the volume $d(k)$ to the corresponding weight. Each demand is further defined by its origin terminal $O(k) \in \mathcal{N}^{\text{ph}}$, destination terminal $D(k) \in \mathcal{N}^{\text{ph}}$, release time $\alpha(k)$ at the origin terminal, and due time $\beta(k)$ at the destination terminal. Shippers are assumed to request either standard or express delivery, indicated by $class(k)$, regardless of their category. Obviously, express deliveries are charged a higher fare than standard deliveries. The parameter $\phi(k)$ specifies the unit fare for each demand, corresponding to its fare class $class(k)$. Let Γ denote the set of container types, with $\gamma(k) \in \Gamma$ representing the container type for each demand. The type of container affects the handling, storage, and transportation costs, due to differing requirements and handling complexities.

Service Network. Let $\mathcal{V}$ denote the set of vessel types, and let F_v represent the maximum number of vessels of type $v \in \mathcal{V}$ available. Each vessel $v \in \mathcal{V}$, pre-assigned to a service, follows a circular sequence of services, starting and ending at the same port. Vessels are characterized by their nominal speed and nominal capacity, measured as $\text{cap}_w(v)$ in tonnage and $\text{cap}_{vol}(v)$ in TEUs. Additional characteristics include length $\text{len}(v)$, maximum draught $dh^+(v)$, minimum draught $dh^-(v)$ and the load-draught coefficient $\theta(v)$, measured in meters per tonnage, which quantifies how changes in loaded weight affect the vessel's draught. The set of potential services

the carrier may operate to meet transportation demand is denoted by Σ. Each potential service $\sigma \in \Sigma$ is defined by its route in the physical network, by its schedule, and by its vessel type. The route of a potential service σ is specified as an ordered set of consecutive stops, including the origin, destination, and intermediate stops, denoted as $\mathcal{N}^{\text{ph}}(\sigma) = \{\eta_i(\sigma) \mid i = 0, \ldots, n(\sigma)\}$. Here, $n(\sigma) = |\mathcal{N}^{\text{ph}}(\sigma)| - 1$ and i indicates the i^{th} stop of the service, with $\eta_0(\sigma) = O(\sigma)$ and $\eta_n(\sigma) = D(\sigma)$ as the origin and destination of the service, respectively. These services are characterized by schedules that specify the departure and arrival times, $\alpha(\eta_i(\sigma))$ and $\beta(\eta_i(\sigma))$, respectively, at each terminal $\eta_i(\sigma)$ in $\mathcal{N}^{\text{ph}}(\sigma)$. Each service has a total duration $\delta(\sigma)$, which includes the time spent at each stop as well as the travel time for each leg. A leg $l_i(\sigma)$ is defined as the segment between each pair of consecutive stops and is expressed as $(\eta_{i-1}(\sigma), \eta_i(\sigma))$ for $i = 1, \ldots, n(\sigma)$.

The WL-SSND-RRM problem is formulated on a time-space network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$. This network is based on time discretization over the schedule length T, divided into equal-length time periods, $t \in \{0, 1, \ldots, T-1\}$. The node set $\mathcal{N}$ is defined as $\{(\eta, t) \mid \eta \in \mathcal{N}^{\text{ph}}, t = 0, \ldots, T-1\}$, representing all terminals in the physical network at each time instant. The arc set $\mathcal{A}$ is the union of moving arcs and holding arcs, $\mathcal{A} = \mathcal{A}^M \cup \mathcal{A}^H$. The set $\mathcal{A}^M$ represents movements between nodes and is defined as: $\mathcal{A}^M = \{((\eta, t), (\eta', t')) \mid \eta, \eta' \in \mathcal{N}^{\text{ph}}, t, t' \in \{0, \ldots, T-1\}, t < t'\}$ This indicates movements between nodes η and η', departing at time t and arriving at time t'. The set $\mathcal{A}^H$ is defined as: $\mathcal{A}^H = \{((\eta, t), (\eta, t+1)) \mid \eta \in \mathcal{N}^{\text{ph}}, t \in \{0, \ldots, T-1\}\}$. These represent a one-time period waiting at terminal η at time t for vessels, demand, and services. According to the definition of moving arcs in the time-space network, each service leg corresponds to a moving arc. Specifically, a moving arc standing for service leg $l_i(\sigma) = \{(\eta_{i-1}(\sigma), \eta_i(\sigma)) \mid i = 1, \ldots, n(\sigma), \sigma \in \Sigma\}$, is defined as $a_{l_i(\sigma)} = ((\eta_{i-1}(\sigma), \alpha(\eta_{i-1}(\sigma))), (\eta_i(\sigma), \beta(\eta_i(\sigma)))$. This arc indicates the departure of the service leg from terminal $\eta_{i-1}(\sigma)$ at time $\alpha(\eta_{i-1}(\sigma))$ and its arrival at terminal $\eta_i(\sigma)$ at time $\beta(\eta_i(\sigma))$. The second type of arc referred to as a holding arc, is thus defined as $a_{\eta t} = ((\eta, t), (\eta, t+1))$, where $(\eta, t) \in \mathcal{N}$. As we formulated the problem on a time-space network, the characteristics of nodes and arcs in the physical network are represented in the time-space network. The water levels, which are defined for each physical link over different periods, can be directly mapped onto moving arcs in the time-space network. Specifically, the water level for a physical arc, denoted as p_a, is represented as $p_{a_{l_i(\sigma)}}$ for moving arcs $a_{l_i(\sigma)} \in \mathcal{A}^M$. Similarly, the berthing capacity Q_η for nodes is represented as $Q_{a_{\eta t}}$ for holding arcs $a_{\eta t} \in \mathcal{A}^H$.

In this setting, we consider the following information revelation and decision-making process. At the tactical planning stage, a transportation plan is developed that encompasses the establishment of a network of scheduled services operated by assigned vessels and the determination of itineraries for regular and profitable spot shippers' demand. Since each service is directly linked to a vessel, the planning process implicitly manages vessel scheduling as part of resource management. However, it also explicitly includes additional resource management decisions, such as determining the number of temporarily idle vessels awaiting their next service departure and the total number of vessels utilized in the service plan. These decisions—encompassing the establishment of a network of scheduled services, the determination of demand itineraries, and explicit resource management—are made during the planning stage, before observing actual water level realizations, and are based solely on statistical distributions of water levels. Once these decisions are made and just before service operations begin (i.e., at the start of the schedule length), we assume that more precise water level forecasts can be obtained, poten-

tially due to specialists' ability to translate updated short-term weather and hydrological data covering the schedule length into accurate water level predictions. These water level forecasts are highly accurate because the schedule length is sufficiently short, ensuring that predictions maintain precision and accurately reflect patterns of weather and hydrological behavior. Thus, improving the accuracy of water level estimations shortly before transportation services begin is both feasible and highly advantageous, as it enables carriers to incorporate the latest data into their decision-making processes, refine tactical plans, and adjust them based on updated water level information. This ensures better alignment with operational realities and helps address limitations caused by fluctuating water levels, which can lead to partial or total reductions in vessel capacity, directly affecting the feasibility of demand itineraries by restricting the amount of demand that can be transported over planned services. Specifically, carriers might re-plan demand itineraries by modifying itineraries or selectively mitigating demand over planned demand itineraries by paying penalties. These adjustments ensure that the transportation plan remains feasible, efficient, and adaptable to dynamic conditions.

Costs. The following unit costs are defined: Let $h(\eta, v)$ denote the holding cost associated with each vessel type at the terminal for one time period, and $\mu(v)$ the cost associated with each vessel type that is not used in the optimal plan. Let $f(\sigma)$ denote the fixed cost of setting up and operating the service σ, and let $c_i(\gamma(k), v(\sigma))$ represent the transportation cost of a container of type $\gamma(k)$ by a vessel of type $v(\sigma)$ on the i^{th} leg of service σ. Here, $\gamma(k) \in \Gamma$ denotes the container type for demand d, and $v(\sigma) \in \mathcal{V}$ specifies the type of the vessel assigned to service σ. Let $c(\eta, \gamma(k))$ denote the cost of holding a container of type $\gamma(k)$ at terminal η for one period and $\kappa(\eta_i(\sigma), \gamma(k))$ represent loading/unloading a container of type$\gamma(k)$ on the i^{th} leg of service σ. Finally, $b(k)$ represents the penalty cost incurred when failing to meet demand $k \in \mathcal{K}$.

4.2 Two-stage Stochastic WL-SSND-RRM Model

Building on the previously described information revelation and decision-making process, we adopt a two-stage stochastic programming framework to establish a tactical plan that maximizes the carrier's overall profit. This tactical plan—including selecting services, managing resources, and determining itineraries for both regular and selected spot shippers—is enriched by the expected adjustment (recourse) costs incurred at the operational level in response to realized water-level conditions, which impact vessel-supported service capacities and may compromise the feasibility of the original demand itineraries. These recourse costs depend not only on the specific water level realizations and their resulting impact on capacity availability but also on the type of adjustment strategy implemented. We define two recourse strategies: (i) *Selective Demand Mitigation Strategy*, in which lower-priority or less profitable demands are selectively refused, prioritizing high-priority shipments even if penalties are incurred for unfulfilled demands. The objective of this recourse strategy is to mitigate the impact of capacity constraints on critical shipments while balancing the penalties associated with unmet demands; and (ii) *Demand Flow Adjustment Strategy*, which focuses on modifying demand itineraries and aims to minimize rerouting and adjustment costs while ensuring demand fulfillment across the service network operated by vessels. These strategies form the foundation for two distinct two-stage stochastic models: *Two-Stage Stochastic Programming with Selective Demand Mitigation Recourse* (2-SPSDM) and *Two-Stage Stochastic Programming with Demand Flow Adjustment*

Recourse (2-SPDFA). In the former, the first-stage tactical plan includes selecting services, managing resources, and determining itineraries for both regular and selected spot shippers, while the second-stage decision involves selectively refusing demands in response to realized water-level conditions that affect vessel-supported service capacities. In the latter, the first-stage decisions are similar, but the second-stage decision focuses on adjusting demand itineraries in response to updated capacity conditions. The objective of each model is to maximize the carrier's overall profit across the two stages by considering the costs associated with tactica level decisions in the first stage and the expected adjustment (recourse) costs in the second stage. The recourse costs in the 2-SPSDM model consist of penalty costs for unmet demand, whereas in the 2-SPDFA model, they encompass adjustment costs for demand itineraries, including demand routing costs (e.g., moving, loading, and unloading) and demand holding costs.

In these stochastic models, a finite scenario set S is used to approximate the underlying probability space, where each scenario $s \in S$ represents a possible realization of water levels on each service leg. Each scenario is assigned a probability ρ^s, with $\sum_{s \in S} \rho^s = 1$. The detailed formulations are presented in the following sections. In this regard, we emphasize that the deterministic model, included in Appendix A.1 and represented by equations (1) to (23), assumes constant water levels throughout the planning horizon. To enhance clarity and facilitate understanding, the deterministic formulation is provided as a reference model, serving as the foundation for developing the two-stage stochastic mathematical models presented in this work.

4.2.1 The 2-SPSDM model

To formulate the 2-SPSDM model, we define five sets of decision variables. The first four sets correspond to the first stage, while the fifth set belongs to the second stage. They are defined as follows:

- **Service Selection ($y(\sigma) \in \{0,1\}$):** Equals 1 if transportation service $\sigma \in \Sigma$ is selected in the tactical plan, and 0 otherwise.
- **Shipper Selection ($\xi(k) \in [0,1]$, $\zeta(k) \in \{0,1\}$):** $\xi(k)$ represents the proportion of the volume of partial-spot demand $k \in \mathcal{K}^{\mathrm{P}}$ to be serviced; $\zeta(k)$ equals 1 if full-spot demand $k \in \mathcal{K}^{\mathrm{F}}$ is accepted, and 0 otherwise.
- **Resource Management ($z(v, a_{\eta_t}) \in \mathbb{Z}_{\geq} 0$, $B(v) \in \mathbb{Z}_{\geq} 0$):** $z(v, a_{\eta_t})$ denotes the number of idle vessels of type v waiting on holding arc a_{η_t} for the departure of their next service; $B(v)$ is the total number of vessels of type v used in the plan.
- **Demand Flow Distribution ($x(k, a_{l_i}(\sigma)) \geq 0$, $\mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma)) \geq 0$, $\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma)) \geq 0$, $\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t}) \geq 0$):** $x(k, a_{l_i}(\sigma))$, $\mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))$, and $\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))$ represent the volumes of demand $k \in \mathcal{K}$ transported, unloaded, and loaded, respectively, on the i^{th} leg of service σ; $\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t})$ is the volume held at terminal η during the time period $(t, t+1)$.
- **Recourse Decisions ($\lambda(k, a_{l_i}(\sigma))^s \geq 0$):** Represents the volume of demand $k \in \mathcal{K}$ that cannot be transported on leg i of service σ water-level realizations under scenario $s \in S$.

Mathematical formulation. The problem is formulated as follows:

$$\begin{aligned}\max \quad & \sum_{k\in\mathcal{K}^{\mathrm{R}}} \phi(k)d(k) + \sum_{k\in\mathcal{K}^{\mathrm{P}}} \phi(k)\xi(k)d(k) + \sum_{k\in\mathcal{K}^{\mathrm{F}}} \phi(k)\zeta(k)d(k) \\ & - \sum_{v\in\mathcal{V}} \mu(v)(F_v - B(v)) - \sum_{\sigma\in\Sigma} f(\sigma)y(\sigma) - \sum_{a_{\eta_t}\in\mathcal{A}^H} \sum_{v\in\mathcal{V}} h(\eta, v)z(v, a_{\eta_t}) \\ & - \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M} \sum_{k\in\mathcal{K}} c_i(\gamma(k), v(\sigma))x(k, a_{l_i}(\sigma)) - \sum_{a_{\eta_t}\in\mathcal{A}^H} \sum_{k\in\mathcal{K}} c(\eta, \gamma(k))\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t}) \\ & - \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M} \sum_{k\in\mathcal{K}} \kappa(\eta_i(\sigma), \gamma(k)) \left(\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma)) + \mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))\right) - \sum_{s\in S} \rho^s G^s(\boldsymbol{\chi})\end{aligned} \tag{26}$$

Subject to

$$(2), (3), (4), (5), (6), (7), (8), (9), (11), (12), (13), (14), (15) - (23)$$

where $G^s(\boldsymbol{\chi})$ is the optimal value of the second-stage problem, and $\boldsymbol{\chi}$ is the vector of the first-stage decisions.

$$G^s(\boldsymbol{\chi}) = \min\ b(k) \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M} \sum_{k\in\mathcal{K}} \lambda(k, a_{l_i}(\sigma))^s \tag{27}$$

Subject to:

$$\theta(v(\sigma))\,\omega \sum_{k\in\mathcal{K}} (x(k, a_{l_i}(\sigma)) - \lambda(k, a_{l_i}(\sigma))^s) + dh^-(v(\sigma)) \le p'^s_{a_{l_i}(\sigma)}, \quad \forall\sigma\in\Sigma,\ a_{l_i}(\sigma)\in\mathcal{A}^M \tag{28}$$

$$\lambda(k, a_{l_i}(\sigma))^s \le x(k, a_{l_i}(\sigma)), \quad \forall k\in\mathcal{K},\ a_{l_i}(\sigma)\in\mathcal{A}^M \tag{29}$$

$$\lambda(k, a_{l_i}(\sigma))^s \ge 0, \quad \forall k\in\mathcal{K},\ a_{l_i}(\sigma)\in\mathcal{A}^M \tag{30}$$

The 2-SPSDM model, formulated in Eqs. (26)–(30), aims to maximize net profit. The first three terms of the objective function represent the revenue from regular shippers, partial-spot shippers, and full-spot shippers, respectively. The costs incurred in the first stage consist of fixed costs and demand flow distribution costs. Fixed costs include the penalty for unused vessels, the fixed cost of setting up and operating services, and the cost of vessels idling at ports awaiting their next departure. In addition to these fixed costs, there are variable costs associated with demand flow distribution—such as moving, holding, loading, and unloading containers. These costs depend on tactical decisions made before the realization of water levels. The final term of the objective function reflects the expected second-stage costs, specifically penalties for unmet demand across all scenarios, each corresponding to a possible realization of water levels(Eq. (27)). The constraints ensure the feasibility of first-stage decisions, including flow conservation constraints (Eqs. (2), (3), (4)), flow conservation on services (Eqs. (5), (6), (7)), service capacity (Eqs. (8), (9)), fleet size and design balance (Eqs. (11), (12), (13)), and terminal berthing capacity (Eq. (14)). The remaining constraints define the boundaries for first-stage decisions. Constraints (28), (29), and (30) ensure the feasibility of second-stage decisions. Specifically, constraint (28) guarantees that the vessel's draught does not exceed the water level for each scenario $s \in S$. Constraints (29) and (30) define the boundaries of second-stage decisions.

4.2.2 The 2-SPDFA model

To formulate the 2-SPDFA model, we define four sets of decision variables. The first three decision variables correspond to the first stage, while the fourth decision variable belongs to the second stage. They are defined as follows:

- **Service Selection ($y(\sigma) \in \{0,1\}$):** Equals 1 if transportation service $\sigma \in \Sigma$ is selected in the tactical plan, and 0 otherwise.
- **Shipper Selection ($\xi(k) \in [0,1]$, $\zeta(k) \in \{0,1\}$):** $\xi(k)$ represents the proportion of the volume of partial-spot demand $k \in \mathcal{K}^{\mathrm{P}}$ to be serviced; $\zeta(k)$ equals 1 if full-spot demand $k \in \mathcal{K}^{\mathrm{F}}$ is accepted, and 0 otherwise.
- **Resource Management ($z(v, a_{\eta_t}) \in \mathbb{Z}_{\geq} 0$, $B(v) \in \mathbb{Z}_{\geq} 0$):** $z(v, a_{\eta_t})$ denotes the number of idle vessels of type v waiting at holding arc a_{η_t} for the departure of their next service; $B(v)$ is the total number of vessels of type v used in the plan.
- **Recourse Decisions ($x(k, a_{l_i}(\sigma))^s \geq 0$, $\mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))^s \geq 0$, $\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))^s \geq 0$, $\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t})^s \geq 0$):** Determine the demand itinerary based on the observed water level scenario s.

Mathematical formulation. The problem is formulated as follows:

$$\begin{aligned}\max \quad & \sum_{k \in \mathcal{K}^{\mathrm{R}}} \phi(k) d(k) + \sum_{k \in \mathcal{K}^{\mathrm{P}}} \phi(k) \xi(k) d(k) + \sum_{k \in \mathcal{K}^{\mathrm{F}}} \phi(k) \zeta(k) d(k) \\ & - \sum_{v \in \mathcal{V}} \mu(v)(F_v - B(v)) - \sum_{\sigma \in \Sigma} f(\sigma) y(\sigma) - \sum_{a_{\eta_t} \in \mathcal{A}^H} \sum_{v \in \mathcal{V}} h(\eta, v) z(v, a_{\eta_t}) - \sum_{s \in S} \rho^s G^s(\boldsymbol{\chi})\end{aligned} \tag{31}$$

Subject to

$$(11), (12), (13), (14), (15) - (19)$$

where $G^s(\boldsymbol{\chi})$ is the optimal value of the second-stage problem, and $\boldsymbol{\chi}$ is the vector of the first-stage decisions.

$$\begin{aligned}G^s(\boldsymbol{\chi}) = \min \quad & \sum_{a_{l_i}(\sigma) \in \mathcal{A}^M} \sum_{k \in \mathcal{K}} c_i(\gamma(k), v(\sigma))\, x(k, a_{l_i}(\sigma))^s + \sum_{a_{\eta_t} \in \mathcal{A}^H} \sum_{k \in \mathcal{K}} c(\eta, \gamma(k))\, \mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t})^s \\ & + \sum_{a_{l_i}(\sigma) \in \mathcal{A}^M} \sum_{k \in \mathcal{K}} \kappa(\eta_i(\sigma), \gamma(k)) \left(\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))^s + \mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))^s\right)\end{aligned} \tag{32}$$

Subject to

$$\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t})^s + \sum_{a_{l_i}(\sigma) \in \mathcal{A}^M, (\eta_{i-1} = \eta, \alpha(\eta_{i-1}) = t)} \mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))^s = \begin{cases} d(k), & \forall k \in \mathcal{K}^{\mathrm{R}}, \eta = O(k), t = \alpha(k) \\ \xi(k) d(k), & \forall k \in \mathcal{K}^{\mathrm{P}}, \eta = O(k), t = \alpha(k) \\ \zeta(k) d(k), & \forall k \in \mathcal{K}^{\mathrm{F}}, \eta = O(k), t = \alpha(k) \end{cases} \tag{33}$$

$$\sum_{\alpha(k) < t \leq \beta(k)} \sum_{a_{l_i}(\sigma) \in \mathcal{A}^M, (\eta_i = \eta, \beta(\eta_i) = t)} \mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))^s = \begin{cases} d(k), & \forall k \in \mathcal{K}^{\mathrm{R}}, \eta = D(k), \\ \xi(k) d(k), & \forall k \in \mathcal{K}^{\mathrm{P}}, \eta = D(k) \\ \zeta(k) d(k), & \forall k \in \mathcal{K}^{\mathrm{F}}, \eta = D(k) \end{cases} \tag{34}$$

$$\begin{aligned} &\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_{t-1}})^s + \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M(\eta_i=\eta,\beta(\eta_i)=t)} \mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))^s \\ &-\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t})^s - \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M((\eta_{i-1}=\eta,\alpha(\eta_{i-1})=t)} \mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))^s = 0 \\ &\forall(\eta,t) \neq (O(k),\alpha(k))\forall\eta \neq D(k), \forall k \in \mathcal{K} \end{aligned} \tag{35}$$

$$\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))^s - x(k, a_{l_i}(\sigma))^s = 0, \forall a_{l_i}(\sigma) \in \mathcal{A}^M,\ \eta_{i-1}(\sigma) = O(\sigma),\ k \in \mathcal{K} \tag{36}$$

$$x(k, a_{l_i}(\sigma))^s - \mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))^s = 0, \forall a_{l_i}(\sigma) \in \mathcal{A}^M,\ \eta_i(\sigma) = D(\sigma),\ k \in \mathcal{K} \tag{37}$$

$$\begin{aligned} &x(k, a_{l_{i-1}}(\sigma))^s - \mathrm{x}^{\mathrm{out}}(k, a_{l_{i-1}}(\sigma))^s + \mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))^s - x(k, a_{l_i}(\sigma))^s = 0, \\ &\forall\sigma \in \Sigma,\ \eta_{i-1} \neq O(\sigma), \eta_i \neq D(\sigma),\ k \in \mathcal{K} \end{aligned} \tag{38}$$

$$\omega \sum_{k\in\mathcal{K}} x(k, a_{l_i}(\sigma))^s \leq cap_w(v(\sigma))y(\sigma), \quad \forall\sigma \in \Sigma, a_{l_i}(\sigma) \in \mathcal{A}^M \tag{39}$$

$$\sum_{k\in\mathcal{K}} x(k, a_{l_i}(\sigma))^s \leq cap_{vol}(v(\sigma))y(\sigma), \quad \forall\sigma \in \Sigma, a_{l_i}(\sigma) \in \mathcal{A}^M \tag{40}$$

$$\theta(v(\sigma))\omega \sum_{k\in\mathcal{K}} x(k, a_{l_i}(\sigma))^s + dh^-(v(\sigma)) \leq p'^s_{a_{l_i}(\sigma)} \forall\sigma \in \Sigma, a_{l_i}(\sigma) \in \mathcal{A}^M \tag{41}$$

$$x(k, a_{l_i}(\sigma))^s \geq 0 \quad \forall k \in \mathcal{K},\ a_{l_i}(\sigma) \in \mathcal{A}^M \tag{42}$$

$$\mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))^s \geq 0 \quad \forall k \in \mathcal{K},\ a_{l_i}(\sigma) \in \mathcal{A}^M \tag{43}$$

$$\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))^s \geq 0 \quad \forall k \in \mathcal{K},\ a_{l_i}(\sigma) \in \mathcal{A}^M \tag{44}$$

$$\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t})^s \geq 0 \quad \forall k \in \mathcal{K},\ a_{\eta_t} \in \mathcal{A}^H \tag{45}$$

The mathematical formulation is presented in Eqs. (31)-(45). The objective function (Eq. (31)) maximizes profit by considering the revenue from satisfying demand across three categories, represented by the first three terms. From this, we subtract the costs associated with first-stage decisions, such as resource management and service selection, as well as the expected second-stage costs, which are captured in the final term—specifically, the cost of demand flow over the scheduled network design across all scenarios Eq. (32). Constraints (11)-(14) ensure the feasibility of the first-stage decisions, focusing on resource management and service selection, while constraints (15)-(19) define the boundaries for these decisions. Constraints (33)-(41) guarantee the feasibility of the second-stage decisions, including flow conservation constraints (Eqs. (33), (34), (35)), flow conservation on services (Eqs. (36), (37), (38)), service capacity (Eqs. (39), (40)), and (Eq. (41)) guarantees that the vessel's draught does not exceed the water level for each scenario $s \in S$. Constraints (42)-(45) define their respective boundaries. It is important to note that the structure of the constraints is consistent with the deterministic model.

5 Solution Approach

In our proposed models, the effects of water level uncertainty are captured using a set of scenarios that represent how service capacities vary randomly. This scenario set must be representative, providing a sufficiently high level of confidence in capturing the variability of observed

service capacity values. Generating such a comprehensive set may require a large number of scenarios and, consequently, the incorporation of substantial contextual information into the optimization process to achieve a detailed and accurate representation of the uncertainty. Therefore, processing and utilizing this information effectively in decision-making can place significant demands on computational resources. Incorporating all scenarios into stochastic WL-SSND-RRM models often renders them numerically intractable, requiring impractical amounts of time and computational power to solve them. This dual challenge—managing the large volume of information while ensuring computational feasibility—represents significant obstacles to the practical application of these models. To address this general challenge, we adopt the two-step approach introduced by (Hewitt et al., 2022) and apply it to our problem setting. Details of both steps are presented in the following subsections.

5.1 Step 1: Scenario Clustering Analysis

Given scenario set $S = \{s_1, s_2, \ldots, s_n\}$, which represents possible realizations of water levels, the primary objective of this step is to interpret these scenarios in terms of the decisions or plans they induce. This analysis aims to better understand how each scenario influences the decisions to be made. We begin by considering the vector of first-stage decisions $\mathcal{X}$, as introduced in Section 4.1 for the 2-SPSDM model and in Section 4.2 for the 2-SPDFA model. These decisions are constrained by a feasible region $\mathcal{F}_\mathcal{X}$, which includes all possible solutions that satisfy the model's constraints, as specified in their respective sections. To assess the performance of these decisions, we solve a restricted single-scenario version of the 2-SPSDM and 2-SPDFA models, using the objective function, g, where $g : \mathcal{F}_\mathcal{X} \times S \to \mathbb{R}^+$. In this context, $g(\chi, s_j)$ represents the objective value (total profit) for a given combination of first-stage decisions $\chi \in \mathcal{F}_\mathcal{X}$ and scenario $s_j \in S$. Among all feasible first-stage decisions in $\mathcal{F}_\mathcal{X}$, the optimal first-stage decision for a specific scenario s_j is denoted by $\chi^*_{s_j}$ and is defined as: $\chi^*_{s_j} = \arg\max_{\chi \in \mathcal{F}_\mathcal{X}} g(\chi, s_j)$. The opportunity cost is then defined as $\theta(s_i \mid s_j) = g(\chi^*_{s_j}; s_j) - g(\chi^*_{s_i}; s_j)$, where $\theta(s_i \mid s_j)$ measures the dissimilarity between decisions optimized for scenario s_i when evaluated under scenario s_j. Since each scenario s_i is compared against every other scenario s_j ($j \neq i$), the *opportunity cost matrix* Θ is formed, where each entry $\theta(s_i \mid s_j)$ reflects the relative performance of the decisions across different scenarios. However, the opportunity cost can be *asymmetric* ($\theta(s_i \mid s_j) \neq \theta(s_j \mid s_i)$), and this asymmetry can introduce biases in the clustering process. To address this, we construct a *symmetric dissimilarity matrix* D by summing the opportunity costs: $D(s_i, s_j) = \theta(s_i \mid s_j) + \theta(s_j \mid s_i)$. This symmetric matrix provides a balanced and mutual measure of dissimilarity between any two scenarios, ensuring its suitability for clustering. We use the symmetric matrix D and apply the K-means Spectral Clustering method, as described in (Hewitt et al., 2022), to form clusters of scenarios based on their decision impact. This process can be approached in two ways: one that seeks to minimize dissimilarity among grouped scenarios and another that aims to maximize dissimilarity. To clearly distinguish between these two approaches, we introduce the following notations.

- Let $C^{\min}_k$, $k = 1, \ldots, K^{\min}$, represent the set of clusters obtained by minimizing dissimilarity. This approach allows us to form clusters of scenarios that lead to similar decisions, thereby reducing redundancy in the scenario set and simplifying it while minimizing information loss.

- Let $C_k^{\max},\ k = 1, \ldots, K^{\max}$, represent the set of clusters obtained by maximizing dissimilarity. This approach enables us to cluster scenarios that result in different decisions, ensuring that each cluster captures diverse, decision-relevant information, thus preserving the richness of uncertainty.

Notably, the number of clusters is determined using the Silhouette Score (Rousseeuw, 1987), a widely used metric for assessing clustering quality.

5.2 Step 2: Bounds Computation

In this section, we present the second step of the solution approach, which focuses on leveraging the scenario clusters obtained in the previous step to apply a series of bounding techniques. These techniques, based on the use of scenario clusters, provide both:

1. A set of alternative feasible solutions, whose objective function values define alternative lower bounds for the stochastic models under consideration.
2. An upper bound value, which serves as a bounding certificate for evaluating the optimality gaps of the obtained feasible solutions.

The *Medoid Lower Bound* is derived using $C_k^{\min},\ k = 1, \ldots, K^{\min}$. This approach is based on the idea of using clusters to reduce the size of the scenario set utilized in the optimization process, thereby retaining a representative subset of scenarios while keeping computational complexity manageable. This representative subset, denoted as M, consists of the *medoid* of each cluster, denoted as M_k. The medoid M_k is defined as the scenario within each cluster that has the smallest average dissimilarity to all other scenarios in that cluster. The probability assigned to each medoid, denoted as p_{M_k}, is calculated as the sum of the probabilities of all scenarios within the same cluster. Formally, if ρ^{s_i} represents the probability of scenario s_i, then the probability of the medoid for cluster $C_k^{\min}$ is given by $p_{M_k} = \sum_{s_i \in C_k^{\min}} \rho^{s_i}$. Given the representative scenarios (medoids) and their corresponding probabilities, the optimal decision $\bar{\chi}$ for the 2-SPSDM and 2-SPDFA models is then determined by maximizing the expected profit over the representative scenario set. Formally, this is given by $\bar{\chi} = \arg\max_{\chi \in \mathcal{F}_\mathcal{X}} \sum_{M_k \in M} g(\chi, M_k) \times p_{M_k}$. The lower bound is then computed by evaluating this solution over all scenarios, expressed as $\sum_{s_i \in S} g(\bar{\chi}; s_i)\rho^{s_i}$, where $g(\bar{\chi}; s_i)$ represents the objective value for scenario $s_i \in S$ given $\bar{\chi}$.

In contrast, the other two bounds, namely the *Cluster Lower Bound* and the *Cluster Upper Bound*, are derived using clusters $C_k^{\max},\ k = 1, \ldots, K^{\max}$. For these bounds, the clustering process partitions the original scenario set in such a way as to produce a representative subset within each cluster that preserves the overall structure of the scenario set. For each cluster $C_k^{\max}$, we define the total probability $P_{C_k^{\max}}$ as: $P_{C_k^{\max}} = \sum_{s_i \in C_k^{\max}} \rho^{s_i}$. The optimal decision $\bar{\chi}_{C_k^{\max}}$ for each cluster is then determined by maximizing the expected profit over the scenarios within that cluster, with scenario probabilities normalized within the cluster. Formally, this is given by: $\bar{\chi}_{C_k^{\max}} = \arg\max_{\chi \in \mathcal{F}_\mathcal{X}} \sum_{s_i \in C_k^{\max}} g(\chi, s_i) \times \frac{\rho^{s_i}}{P_{C_k^{\max}}}$. The upper bound is calculated by aggregating the results from all clusters and is represented as $\sum_{k=1}^{K^{\max}} \sum_{s_i \in C_k^{\max}} g(\bar{\chi}_{C_k^{\max}}; s_i) \times \rho^{s_i}$. The lower bound is derived by comparing results across clusters and is represented as $\min_{k=1,\ldots,K^{\max}} \sum_{s_i \in C_k^{\max}} g(\bar{\chi}_{C_k^{\max}}; s_i) \times \rho^{s_i}$.

6 Numerical experiments

In this section, we present a set of computational experiments and analyses. These serve two main objectives. The first is to assess the performance of the proposed solution approach and validate its ability to balance computational efficiency and solution quality. The second objective is to explore the impact of explicitly considering water level uncertainty on the tactical planning process. We begin by evaluating the performance of the solution approach in Section 6.1. We also investigate several critical managerial questions in the remainder of the section. Section 6.2 examines how incorporating the effects of stochastic water levels within the tactical planning process can improve upon the traditional approach of relying on predictions and applying a deterministic optimization model based on those predictions. In this study, we further propose two recourse strategies to account for the impact of reacting to randomly changing water levels. These strategies both aim to generate more profitable tactical plans (i.e., plans that can lead to higher profits for carriers by better adapting to uncertainty) yet they may result in plans with different structural characteristics. Section6.3 assesses the effectiveness of these alternative recourse strategies. Finally, Section 6.4 investigates the impact of water level fluctuations on metrics such as the number of vessels used, consolidation, cost, and profit.

All implementations are conducted using the Pyomo software package and the Gurobi solver on a machine equipped with an Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz and 256 GB of memory. The test instances are generated for a linear network topology commonly observed in Canadian and European inland waterway systems, where major ports are connected sequentially. Representative examples include the Saint Lawrence River corridor—linking ports such as Montréal, Québec City, Sept-Îles, and Port-Cartier—and the Rhine River system, comprising port sequences along the Upper, Middle, and Lower Rhine sections. The linear network used in our experiments includes four connected terminals. Water levels are considered at critical points along the waterways and range between 150 cm and 350 cm, reflecting observed variability as reported by (Christodoulou et al., 2020). Two vessel types (small and large) are considered, with large vessels incurring twice the operating cost of small ones. A total of 360 potential services are generated across 20 routes, 9 distinct start times, and 2 vessel types within a 14-period weekly cycle. Five demand instances are constructed, each comprising 55 to 66 randomly generated requests with volumes between 5 and 25 TEUs, resulting in a total of 298 demands. The full instance generation process is detailed in Appendix A.2.

6.1 Performance of the solution approach

This section presents a performance analysis of the proposed solution approach. Computational tests are conducted to compare the performance of the proposed approach against Gurobi in solving stochastic models for instance categories labeled Instance-1 through Instance-7, as described in Appendix A.2. To perform the experimental analysis, we follow the solution approach described in Section 5, which includes two steps. Step 1 involves constructing the dissimilarity matrix and conducting the clustering analysis. Step 2 consists of calculating the different bounds using both the medoid-based approach and the cluster-based approach. The total computational time required to perform all steps of the solution approach is recorded. This includes the time needed to compute the dissimilarity matrix, conduct the clustering

analysis, and calculate the different bounds (including the medoid lower bound as well as both the cluster lower and upper bounds). Preliminary testing showed that the 2-SPSDM instances are comparatively easier to solve than those for the 2-SPDFA model. Specifically, the cluster-based approach was extremely fast when applied to the 2-SPSDM model but was slower compared to the medoid-based approach. Therefore, to properly assess the efficiency of the solution approach in this case, we set the maximum time limit for Gurobi to the average computational time taken by the cluster-based approach, ensuring a fair comparison where both solution approaches are allotted comparable computational times. For the more complex 2-SPDFA model, Gurobi is provided with a significantly higher computational time limit (7200 seconds), which is substantially greater than the time required for the cluster-based approach. This ensures that Gurobi has enough time to fully explore the solution space, allowing for a meaningful performance comparison between the two approaches.

Table 1 and Table 2 present the performance evaluation of the proposed solution approach on the 2-SPSDM and 2-SPDFA models, respectively, across instance categories labeled Instance-1 through Instance-7. The "Gap (%)" columns show the average optimality gaps (in percentage) achieved by Gurobi, the Medoid method, and the Cluster method over the three sub-instances associated with each main instance, indicating how closely each method approximates the optimal solution. The optimality gap for the cluster-based approach is calculated as the relative difference between the cluster upper bound and the cluster lower bound obtained through this method. For the medoid-based approach, the optimality gap is determined by calculating the relative difference between the medoid lower bound and the cluster upper bound. The "Time (s)" column under the Cluster method represents the average computational time (in seconds) required to apply the cluster method to solve each problem instance. Rows labeled "Mean," "Median," "Std Dev," "Max," and "Min" summarize the performance metrics across all instances, providing insights into the variability and efficiency of each method. The performance

Table 1: Performance Evaluation of the Proposed Solution Approach on the 2-SPSDM Model

Instance	Gurobi	Medoid	Cluster	
	Gap (%)	Gap (%)	Gap (%)	Time (s)
Instance-1	4.99	2.67	1.71	19.97
Instance-2	17.3	0.51	1.56	22.05
Instance-3	2.74	5.29	2.74	25.41
Instance-4	3.77	7.34	2.44	23.81
Instance-5	4.72	5.15	0.97	21.15
Instance-6	10.1	9.50	2.73	20.23
Instance-7	5.74	1.22	2.02	20.74
Mean	7.05	4.38	2.03	21.34
Median	4.99	5.15	2.02	21.15
Std Dev	5.18	3.14	0.60	2.07
Max	17.30	9.50	2.74	25.41
Min	2.74	0.51	0.97	19.97

evaluation for the 2-SPSDM model (Table 1) indicates that the Cluster method consistently outperforms Gurobi in terms of optimality gap. The average gap for the Cluster method is 2.03%, significantly lower than Gurobi's average gap of 7.05%, highlighting the efficiency of the Cluster method. The maximum gap observed for the Cluster method is 2.74%, whereas Gurobi reaches a much higher maximum gap of 17.30%. The minimum gap for the Cluster method is 0.97% in Instance-5, with an average computation time of 21.34 seconds. Notably, in the same computation time, Gurobi achieves a gap of 4.72% for Instance-5, indicating a

substantial improvement by the Cluster method. Importantly, when the time constraint is removed, Gurobi solves to optimality in an average of 98.57 seconds, compared to 21.34 seconds for the Cluster method, indicating the latter's superior efficiency under limited-time conditions. The Medoid method also demonstrates improved performance over Gurobi in some instances, achieving lower optimality gaps in cases such as Instance-2, where the gap is reduced to 0.51%. However, the Medoid method shows greater variability across instances, with a mean gap of 4.38%, a maximum gap of 9.50% (Instance-6), and a minimum gap of 0.51% (Instance-2). This variability suggests that while the Medoid method can be effective, its performance is less consistent compared to the Cluster method. In particular, the medoid bound is based on applying scenario reduction, which effectively eliminates information when solving the optimization problem. This reduction simplifies the problem but also introduces a trade-off, as removing scenarios inherently limits the model's ability to fully capture the variability in the uncertainty representation. In contrast, the cluster bounds leverage the full original scenario set, preserving all information contained within it.

Table 2: Performance Evaluation of the Proposed Solution Approach on the 2-SPDFA Model

Instance	Gurobi	Medoid	Cluster	
	Gap (%)	Gap (%)	Gap (%)	Time (s)
Instance-1	5.88	7.53	0.19	140.05
Instance-2	11.1	6.70	0.30	104.55
Instance-3	5.43	2.83	1.19	300
Instance-4	42.9	18.53	11.54	255.88
Instance-5	3.03	6.40	1.24	399.58
Instance-6	9.38	10.50	0.53	203.65
Instance-7	6.25	4.22	0.01	268.70
Mean	11.28	8.24	2.29	238.21
Median	6.25	6.70	0.53	255.88
Std Dev	13.53	5.32	4.11	109.91
Max	42.90	18.53	11.54	399.58
Min	3.03	2.83	0.01	104.55

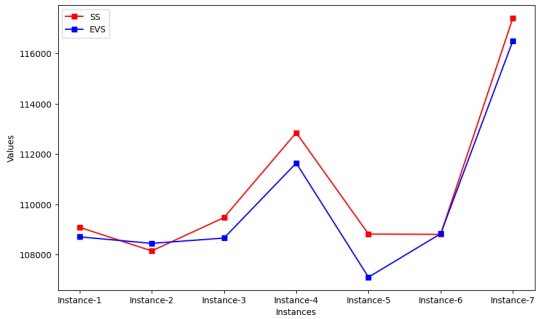
Table 2 presents the performance evaluation of the proposed solution approach on the 2-SPDFA model across multiple instances, comparing optimality gaps and computation times for the Medoid and Cluster methods, with Gurobi set to a fixed runtime of 7200 seconds. The Cluster method achieves the lowest average gap at 2.29%, significantly outperforming Gurobi, which has an average gap of 11.28%. Additionally, the minimum and maximum gaps for the Cluster method are 0.01% (Instance-7) and 11.54% (Instance-4), respectively, while Gurobi's minimum and maximum gaps are 3.03% (Instance-5) and 42.9% (Instance-4). The narrower range of gaps for the Cluster method, compared to the wider range for Gurobi, indicates highly effective performance across instances. However, the average computation time for the Cluster method is 238.21 seconds, compared to Gurobi's fixed runtime of 7200 seconds. The Medoid method demonstrates significant performance improvements over Gurobi, reducing the average gap from 11.28% to 8.24% and the standard deviation from 13.53 to 5.32, indicating not only better average gaps but also significantly reducing the variability of results.

In summary, Tables 1 and 2 show that both the Medoid and Cluster methods significantly outperform the direct use of Gurobi for solving the stochastic formulations, achieving substantial reductions in gaps. Additionally, the average computation times reported in Tables 1 and 2 indicate that the Cluster method is an efficient approach for successfully tackling large-scale

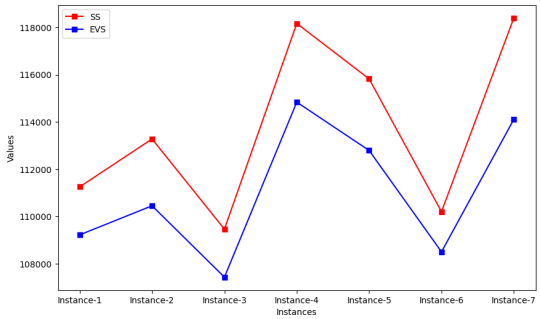
instances of the considered models.

6.2 Importance of modeling uncertainty

The Value of Stochastic Solution (VSS) is a measure used in stochastic programming to quantify the overall benefit of using a stochastic model over a deterministic model when making decisions under uncertainty. VSS measures the expected improvement in performance—here profit maximization—achieved by basing decisions on a stochastic model rather than a deterministic approximation that instead uses a prediction for the stochastic parameters.



(a) Comparison of SS and EVS for 2-SPSDM model



(b) Comparison of SS and EVS for 2-SPDFA model

Figure 2: Value of the Stochastic Solution

To calculate the VSS, two key values need to be determined. *Expected Value Solution (EVS)*: This is the outcome of applying the solution from a deterministic model that does not account for uncertainty and instead uses expected values of uncertain water levels. *Stochastic Solution (SS)*: This represents the best upper bound obtained using the proposed solution approach for the stochastic model, which incorporates a set of possible scenarios and their associated probabilities for water levels. Figure 2a and Figure 2b present the VSS results for the 2-SPSDM and 2-SPDFA models, respectively. In these figures, the blue line with square markers represents the EVS, derived from a deterministic model that uses average water levels without accounting for uncertainty. The red line with square markers represents the SS, which explicitly accounts for random fluctuations in water levels.

Figure 2a, we observe that, in most instances (except for Instance 2), the SS performs better or similarly to the EVS. In Instance 2, the SS performs marginally worse because the stochastic model is not solved to optimality for one of the three instances included in Instance 2. In Instance 6, both models exhibit nearly identical performance, while in Instances 4, 5, and 7, the SS significantly outperforms the EVS, with the greatest improvement seen in Instance 5. This pattern highlights the importance of solving the 2-SPSDM stochastic model when the infeasibilities in the tactical plan due to varying water levels are penalized by applying a cost to the amount of demand that cannot be transported. A similar pattern is observed in Figure 2b, which compares the SS and EVS values for the 2-SPDFA model. As shown, the SS consistently outperforms the EVS across all instances, underscoring the effectiveness of the 2-SPDFA model in providing a tactical plan that is better suited to deal with uncertainty in water levels.

6.3 Model Evaluation

This section includes a set of comparative tests designed to evaluate the performance of the proposed models for the WL-SSND-RRM problem in a stochastic environment. The objective is to assess the models' ability to generate profitable tactical plans and to analyze how different recourse strategies lead to plans with distinct structural characteristics. To provide a more accurate basis for evaluating the obtained solutions with a higher level of confidence as to their actual value (the so-called ground truth), we define a Ground Truth set consisting of 500 scenarios sampled from the same distribution described in Appendix A.2. Each test begins by independently solving each model to obtain its tactical plan. Subsequently, the performance of each solution is evaluated by applying it either to a different model or within its own framework using the Ground Truth scenarios. To facilitate this evaluation, we define two types of benchmarks: Cross-Model Benchmark and Internal Model Benchmark:

- **Cross-Model Benchmark:** This benchmark includes the Deterministic-in-2-SPSDM, Deterministic-in-2-SPDFA, and 2-SPSDM-in-2-SPDFA frameworks. These frameworks involve fixing certain decisions derived from one model (e.g., the deterministic solution or 2-SPSDM) as input to another stochastic model framework (e.g., 2-SPSDM or 2-SPDFA). The resulting restricted stochastic models are then solved by exploring the remaining decision subspace to complete the full solution using the ground truth scenarios.
- **Internal Model Benchmark:** This benchmark includes the 2-SPSDM Baseline and 2-SPDFA Baseline frameworks. In these frameworks, each model's optimized solution is evaluated within its original framework using the ground truth scenarios, establishing a baseline for that model's optimal performance under its specific recourse strategy.

Based on these benchmarks, four comparative tests are performed. The first two tests, $\Delta 1$: Deterministic-in-2-SPSDM vs. 2-SPSDM Baseline and $\Delta 2$: Deterministic-in-2-SPDFA vs. 2-SPDFA Baseline, are designed to validate the VSS analyses previously conducted by performing out-of-sample assessments for the two stochastic models. The third test, $\Delta 3$: 2-SPSDM-in-2-SPDFA vs. 2-SPDFA Baseline, assesses whether a tactical plan obtained using a model that applies itinerary infeasibility penalty costs can serve as a good proxy for a plan obtained using a model that applies the more involved and complex flow optimization recourse costs. Specifically, this test seeks to determine if there is additional value, from a tactical planning perspective, in solving the more complex 2-SPDFA stochastic model. The final test, $\Delta 4$: Deterministic-in-2-SPSDM vs. Deterministic-in-2-SPDFA, compares the effectiveness of two stochastic models, each defined by a distinct approach to handling the impact of randomly changing water level conditions.

In Table 3, we compare the performance of two stochastic programming models across four tests ($\Delta 1$, $\Delta 2$, $\Delta 3$, and $\Delta 4$), focusing on two key metrics: Relative Difference in Recourse Decision Value (RD) and Relative Difference in Profit Value (RP). These metrics evaluate each model's adaptability and profitability when subjected to uncertainty. The RD column represents the percentage difference in recourse decision values between the two benchmarks, illustrating each model's effectiveness in meeting demand under variable conditions. RD in $\Delta 1$ is computed as the relative difference in the total amount of unmet demand, highlighting how each model manages demand satisfaction under uncertainty. RD in $\Delta 2$ and $\Delta 3$ capture

Table 3: Comparison tests

Test	Benchmark	RD 1 to 2 (%)	RP 1 to 2(%)
$\Delta 1$	1: Deterministic-in-2-SPSDM 2: 2-SPSDM Baseline	47.21	-3.04
$\Delta 2$	1: Deterministic-in-2-SPDFA 2: 2-SPDFA Baseline	-25.23	-2.03
$\Delta 3$	1: 2-SPSDM-in-2-SPDFA 2: 2-SPDFA Baseline	-19.86	1.90
$\Delta 4$	1: Deterministic-in-2-SPSDM 2: Deterministic-in-2-SPDFA	-39.73	-8.81

the relative difference in the total amount of met demand, providing a direct comparison of the models' performance in fulfilling demand. RD in $\Delta 4$ is calculated by first determining the total met demand in the first benchmark as the difference between total demand and unmet demand, and then comparing this value to the total met demand in the second benchmark. The RP column, on the other hand, measures the relative difference in profit between the two benchmarks.

The RD values for Δ_1 and Δ_2 highlight that when service selections and demand itineraries are planned without accounting for variability, these plans become less reliable under fluctuating water conditions. In contrast, the tactical plans obtained using the 2-SPSDM model and the 2-SPDFA model demonstrate better performance in fulfilling demand, with 47.21% less unmet demand and 25.23% more transported demand, respectively. Similarly, the RP values for Δ_1 and Δ_2 highlight the overall benefit of using the two stochastic models over a deterministic model when making decisions under uncertainty, with an improvement in profit of 3.04% and 2.03%, respectively.

$\Delta 3$ highlights the tactical plans obtained using the 2-SPSDM model can nonetheless lead to higher profitability when itineraries are adjusted based on the revealed water levels. This can be explained by the fact that the plans obtained tend to select fewer services. Here, an average profit increase of 1.90% is observed across all results obtained. On the other hand, the plans obtained by solving the 2-SPDFA model include a higher number of services, allowing for more transportation paths when defining itineraries. However, these additional options come with higher fixed costs, which are adverse to profit maximization. Nevertheless, having access to a greater number of potential itineraries significantly decreases the amount of unmet demand, with an observed average reduction of 19.86% in serviced demands when using the 2-SPSDM model instead of the 2-SPDFA model. Second, the difference in profits can be attributed to the fact that the instances are not guaranteed to be solved to optimality. Thus, additional profit-maximizing benefits could still be gained from using the 2-SPDFA model by further improving the quality of the solutions through additional computational efforts. Lastly, this comparison illustrates a general trade-off between prioritizing profit maximization and the amount of serviced demand. Solving the 2-SPSDM model can produce tactical plans conducive to profit maximization, even when such plans are implemented in a different setting where transportation itineraries are allowed to be adjusted based on revealed water levels. However, such plans are not necessarily set up to easily adjust the flows according to varying water level conditions, leading to a reduction in the amount of serviced demand. Considering that there may be significant opportunity losses in not servicing demands, the 2-SPDFA model can provide additional value by leading to plans that accommodate more alternative demand itineraries.

In $\Delta 4$, the goal is to compare the efficiency of different recourse strategies in stochastic models in terms of profit maximization and the amount of serviced demand when dealing with changing water levels. As shown in the table, when the carrier adopts a more flexible recourse strategy, such as demand flow adjustment in the 2-SPDFA model, it achieves a higher profit (8.81%) and can transport significantly more demand (39.73%). These results highlight the substantial benefits of adjusting itineraries in response to water level variations.

6.4 Water Level Impact

In this section, we examine the performance, efficiency, and structural variations of solutions obtained under three different water levels—maximum, mean, and first-quarter—ranging from 150 to 350 cm, as defined by(Christodoulou et al., 2020). These water levels serve as the basis for evaluating how fluctuations impact key metrics such as overall system performance, customer satisfaction, and whether any systematic structural changes emerge as the water levels vary. Our goal is to measure the relationship between water level fluctuations and the level of multi-service usage, a critical factor in barge transportation network design for minimizing operational costs and optimizing resource use. A decrease in the number of demands using multiple paths (services) to reach their destinations typically indicates reduced flexibility and efficiency in the network. Additionally, we investigate the relationship between water levels and the distribution of small and large vessels, as vessel deployment is directly influenced by water depth. Changes in water levels affect vessel draught, which in turn impacts the types of vessels that can be effectively deployed under varying conditions. We also assess the impact of water levels on operational costs, driven by changes in the network structure—specifically, the shifts in multi-service usage and vessel size. Furthermore, we investigate the proportion of spot shippers that are satisfied at each water level, as customer satisfaction directly influences the network's profitability. We also examine capacity usage as a measure of how efficiently the deployed transport capacity is utilized under different water levels. This comprehensive analysis aims to provide a clearer understanding of how water level fluctuations affect the network's performance, cost-efficiency, and financial outcomes.

6.4.1 Performance Indicators (PIs)

To evaluate the experimental results, we use the following performance indicators for each water level (maximum, mean, and first quarter):

Structural Impact PIs

- **Multi-path Usage:** For each solution, we record the number of services used by each demand across all instances. We then produce a histogram to show the distribution of path usage frequencies across all demands. For example, if a demand uses two paths, it is counted as such.
- **Vessel Size:** We record the different sizes of vessels used to transport demands for each solution and produce a histogram showing the distribution of vessel sizes. This allows us to assess how water levels influence the deployment of small and large vessels.

Performance PIs

- **Profit:** A reflection of the overall system's performance, considering revenue obtained and total costs.
- **Demand Flow Costs:** The costs incurred for moving, loading, and unloading demands through the network.
- **Service Costs:** The expenses associated with the transport services used.
- **Shipper Satisfaction (TEUs):** Measured as the percentage of the volume of spot shippers who are fully or partially met under each water level condition.
- **Capacity Usage:** Defined as the ratio of total volume-kilometers moved to total capacity-kilometers operated.

6.4.2 Results and analysis

In the figures presented, we examine multi-path usage and vessel usage at different water levels—350 cm, 250 cm, and 200 cm—highlighting how the transportation network responds to fluctuating water conditions. These figures provide important insights into the operational challenges and adaptations the system may make as water levels change.

Figure 3 illustrates the multi-path usage at the three water levels. At 350 cm, the network satisfies 207 total demands, with 157 of those demands satisfied by a single service and 50 demands requiring multiple services. This suggests that at higher water levels, most demands are transported via one service, showing less need for alternative services. As we move to 250 cm, the total number of satisfied demands decreases slightly to 200, and the distribution of service usage shifts. 127 demands now use a single service, while 73 demands (61 using two services, 12 using three) require more than one service to reach their destination. This increase in multi-path usage reflects the growing operational challenge as the water level drops, forcing the network to rely on more services to maintain service levels. At 200 cm, the system is under more strain, satisfying only 185 demands. The number of single-service demands decreases to 104, while 70 demands require two services, 10 demands need three, and 1 demand uses four services. In total, 81 demands rely on multiple paths, which is a 62% increase compared to the 350 cm level. This shows that as water levels decrease, the network faces greater operational pressure, leading to more fragmented routing and less efficient transportation.

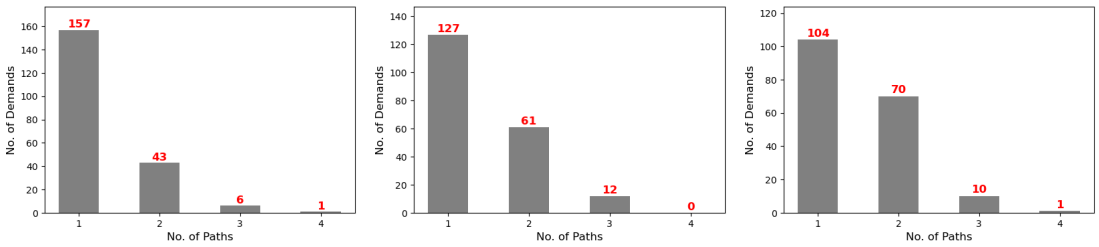


Figure 3: Multi-path usage at different water levels: 350 cm, 250 cm, and 200 cm.

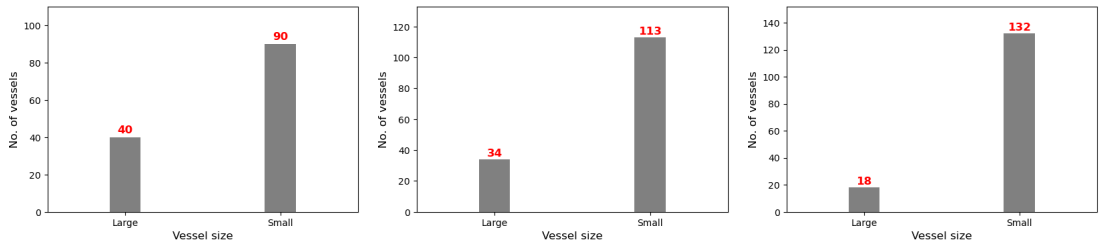


Figure 4: Vessel usage at different water levels: 350 cm, 250 cm, and 200 cm.

Figure 4 highlights the vessel usage at the same water levels. At 350 cm, the network operates with 40 large vessels and 90 small vessels, indicating that 130 services are activated to meet demand. As the water level drops to 250 cm, the number of large vessels decreases to 34, while the number of small vessels increases to 113, resulting in 147 total services being activated. This shift shows that as the water becomes shallower, the system compensates by deploying more small vessels, which are capable of operating under these conditions but require more services due to their limited capacity. At 200 cm, this trend becomes even more pronounced, with only 18 large vessels in operation and a significant increase in the number of small vessels to 132, resulting in a total of 150 activated services—20 more than at the 350 cm water level. This reflects the severe limitations imposed by low water levels on the network's ability to use larger vessels. Small vessels become the primary means of transportation, but their lower cargo capacity necessitates higher numbers of activated services. This leads to increased operational costs (as highlighted in Table 4).

Overall, the figures show that multi-path usage increases as water levels drop, with more demands requiring multiple routes. This indicates a more fragmented and less efficient system at lower water levels. Additionally, vessel usage indicates a clear shift towards smaller vessels as water levels decrease, leading to an increased number of activated services. The total number of services rises from 130 at 350 cm to 150 at 200 cm, reflecting the system's need to compensate for reduced capacity by deploying more smaller vessels. This combination of increased multi-path usage and reliance on smaller vessels at lower water levels results in higher operational costs, reduced profitability, and lower shipper satisfaction, as shown in the table 4.

Table 4: Performance Evaluation of Solutions at Different Water Levels

Water level	Profit	Costs		Shipper Satisfaction (%)		Capacity Usage (%)
		Demand Flow Cost	Service Cost	Partial Spot	Full Spot	
350 cm	568140	134506	11900	62.39	57.61	61.28
250 cm	565725	133927	12670	59.43	53.18	56.31
200 cm	559936	132518	11760	57.00	51.09	51.47

The table reveals noticeable variations in profit, costs, and shipper satisfaction as water levels decrease. At 350 cm, the profit is 568,140. However, as water levels drop to 250 cm, the profit decreases by 0.42% to 565,725, and at 200 cm, it drops further by 1.46% to 559,936. While the absolute change might appear small, the decline is consistent and shows that lower water levels negatively impact profitability. This reduction in profit reflects the system's increasing inefficiency, where higher operational costs are driven by the need to compensate for reduced

capacity by using more small vessels and increasing multi-path usage.This is further evidenced by the decline in capacity usage, which drops from 61.28% at 350 cm to 56.31% at 250 cm, and then to 51.47% at 200 cm. Despite deploying a higher number of vessels and services, the system transports less freight per unit of available capacity, indicating rising inefficiency under lower water levels. The demand flow cost, although it decreases slightly from 134,506 at 350 cm to 132,518 at 200 cm (a 1.48% reduction), masks the real inefficiencies in the system. The marginal reduction in flow cost likely results from fewer demands being met, as indicated by the drop in shipper satisfaction, rather than an improvement in operational efficiency. In contrast, service costs exhibit more noticeable fluctuations. From 11,900 at 350 cm, service costs increase by 6.48% to 12,670 at 250 cm, reflecting the increased use of smaller vessels that are less efficient and require more frequent services. By 200 cm, service costs drop slightly to 11,760 (a 7.18% drop compared to 250 cm). This reduction at 200 cm may indicate reduced activity in the network due to fewer overall demands being satisfied, but it does not imply improved operational efficiency. The drop in shipper satisfaction is stark, particularly in the full spot shippers category, which declines from 57.61% at 350 cm to 51.09% at 200 cm, representing an 11.32% decrease. Similarly, partial spot shipper satisfaction falls by 8.65%, from 62.39% to 57.00%. These declines reflect the system's increasing inability to fully meet demand as water levels decrease. The reduced capacity, due to the shift towards smaller vessels, directly limits the number of shipments that can be satisfied, causing a significant drop in customer satisfaction.

In summary, the drop in water levels from 350 cm to 200 cm leads to a cumulative reduction in profit by 1.46%, driven by increased operational inefficiencies. Although demand flow costs appear stable, the 6.48% spike in service costs at 250 cm highlights the significant inefficiencies that arise as the network compensates by using smaller vessels more frequently. Additionally, the drop in shipper satisfaction—particularly the 11.32% decline in full spot shippers—illustrates the growing challenges in meeting demand. Overall, these results underscore the direct impact of lower water levels on profitability and customer satisfaction, driven by higher service costs and reduced network capacity.

7 Conclusions

In this study, we introduced the Water-Level-Constrained Scheduled Service Network Design with Resource and Revenue Management (WL-SSND-RRM) problem, which addresses the challenges of integrating water-level uncertainty into tactical planning for barge transportation systems. To explicitly account for the effects of water level uncertainty in tactical planning for barge transportation, we propose two alternative stochastic models, each defined by a specific approach to handling the impact of randomly changing water level conditions. The first stochastic model relies on the application of penalties directly tied to the amount of demand that cannot be routed under the established tactical plan. The second one, instead, assumes that demand itineraries can be adjusted based on observed water levels, allowing for flows to dynamically adapt to fluctuating conditions. To address the stochastic models, we apply a general decision-based scenario clustering analysis approach, which enables the computation of a series of alternative bounding techniques. Through numerical experiments, these techniques are shown to significantly outperform a high-performance commercial solver in both compu-

tational efficiency and solution quality. This study also presented a set of comparative tests designed to evaluate the performance of the proposed models. Compared to the deterministic model, these tests highlight the importance of explicitly accounting for uncertainty in tactical planning. Moreover, the results compare the two stochastic models to assess the structural differences introduced in the tactical plan and their varying effectiveness in handling water level uncertainty. Finally, the study highlighted the significant influence of water levels on key tactical decisions (specifically, service selection and demand routing), as well as profitability and operational efficiency.

Several avenues for future research remain open. One direction is the development of novel exact solution methods to solve the proposed stochastic WL-SSND-RRM model. Additionally, future studies could extend the problem by incorporating other sources of uncertainty, such as variability in travel times, to enhance the robustness and applicability of the proposed methodologies. Another promising direction involves exploring alternative recourse strategies. For instance, explicitly managing the vessel fleet could enable dynamic reassignment of vessels to services, allowing the recourse strategy to effectively address infeasibilities caused by random fluctuations in water levels.

Acknowledgments

While working on the project, the first author was Ph.D. candidate at the School of Management, Université du Québec à Montréal (UQAM), Canada, and member of CIRRELT. The third author held the UQAM Chair in Intelligent Logistics and Transportation Systems Planning, and was Adjunct Professor in the Department of Computer Science and Operations Research at the Université de Montréal, while the fourth author held the Canada Research Chair in Stochastic Optimization of Transport and Logistics Systems. We gratefully acknowledge the financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through its Collaborative Research and Development, and Discovery grant programs. We also gratefully acknowledge the support of Fonds de recherche du Québec through their infrastructure grants. Finally, during the preparation of this work the authors used ChatGPT-4, under its most strict privacy enforcing settings, for text edit purposes exclusively. After using this tool, the authors reviewed and further edited the text as needed and take full responsibility for the content of the publication.

References

Agamez-Arias, A.d.M., Moyano-Fuentes, J., 2017. Intermodal transport in freight distribution: a literature review. Transport Reviews 37, 782–807.

Bai, R., Wallace, S.W., Li, J., Chong, A.Y.L., 2014. Stochastic service network design with rerouting. Transportation Research Part B: Methodological 60, 50–65.

Bakir, I., Erera, A., Savelsbergh, M., 2021. Motor carrier service network design, in: Crainic, T., Gendreau, M., Gendron, B. (Eds.), Network Design with Applications in Transportation and Logistics. Springer, Boston, MA. chapter 16, pp. 427–467.

Bektaş, T., Crainic, T.G., 2008. A Brief Overview of Intermodal Transportation, in: Taylor, G.D. (Ed.), Logistics Engineering Handbook. Taylor and Francis Group, Boca Raton, FL, USA. chapter 28, pp. 1–16.

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

Braekers, K., Caris, A., Janssens, G.K., 2013. Optimal shipping routes and vessel size for intermodal barge transport with empty container repositioning. Computers in Industry 64, 155–164.

Caris, A., Macharis, C., Janssens, G.K., 2012. Corridor network design in hinterland transportation systems. Flexible Services and Manufacturing Journal 24, 294–319.

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

Christodoulou, A., Christidis, P., Bisselink, B., 2020. Forecasting the impacts of climate change on inland waterways. Transportation Research Part D: Transport and Environment 82, 102159.

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

Crainic, T., 2003. Long-haul freight transportation, in: Hall, R. (Ed.), Handbook of Transportation Science. second ed.. Kluwer Academic Publishers, Norwell, MA, pp. 451–516.

Crainic, T., Hewitt, M., 2021. Service network design, in: Crainic, T., Gendreau, M., Gendron, B. (Eds.), Network Design with Applications in Transportation and Logistics. Springer, Boston, MA. chapter 12, pp. 347–382.

Crainic, T., Kim, K., 2007. Intermodal transportation, in: Barnhart, C., Laporte, G. (Eds.), Transportation. North-Holland, Amsterdam. volume 14 of *Handbooks in Operations Research and Management Science*. chapter 8, pp. 467–537.

Crainic, T., Rei, W., 2025. 50 Years of Operations Research for Planning Consolidation-based Freight Transportation. Research Report CIRRELT-2024-11. Centre interuniversitaire de

recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Université de Montréal. Montréal. Forthcoming in EURO Journal on Transportation and Logistics.

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

Crainic, T.G., Hewitt, M., Maggioni, F., Rei, W., 2021. Partial benders decomposition: general methodology and application to stochastic network design. Transportation Science 55, 414–435.

Crainic, T.G., Hewitt, M., Rei, W., 2014. Scenario grouping in a progressive hedging-based meta-heuristic for stochastic network design. Computers & Operations Research 43, 90–99.

van Dorsser, C., Vinke, F., Hekkenberg, R., van Koningsveld, M., 2020. The effect of low water on loading capacity of inland ships. European Journal of Transport and Infrastructure Research 20, 47–70.

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

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

Hoff, A., Lium, A.G., Løkketangen, A., Crainic, T.G., 2010. A metaheuristic for stochastic service network design. Journal of Heuristics 16, 653–679.

Jiang, X., Bai, R., Wallace, S.W., Kendall, G., Landa-Silva, D., 2021. Soft clustering-based scenario bundling for a progressive hedging heuristic in stochastic service network design. Computers & Operations Research 128, 105182.

Kaut, M., 2014. A copula-based heuristic for scenario generation. Computational Management Science 11, 503–516.

Kaut, M., Vladimirou, H., Wallace, S.W., Zenios, S.A., 2007. Stability analysis of portfolio management with conditional value-at-risk. Quantitative Finance 7, 397–409.

Lanza, G., Crainic, T., Passacantando, M., Scutellà, M., 2024. Continuous-Time Service Network Design with Stochastic Travel Times. Research Report CIRRELT-2024-29. Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport, Université de Montréal, Université de Montréal. Montréal, QC, Canada.

Lanza, G., Crainic, T.G., Rei, W., Ricciardi, N., 2021. Scheduled service network design with quality targets and stochastic travel times. European Journal of Operational Research 288, 30–46.

Lium, A.G., Crainic, T.G., Wallace, S.W., 2009. A study of demand stochasticity in service network design. Transportation Science 43, 144–157.

Müllerklein, D., Fontaine, P., 2025. Resilient transportation network design with disruption uncertainty and lead times. European Journal of Operational Research 322, 827–840.

Prandtstetter, M., Widhalm, P., Leitner, H., Czege, I., 2023. A novel visualisation tool for reliable inland navigation. Transportation Research Procedia 72, 3537–3544.

Riessen, B.V., Negenborn, R.R., Dekker, R., Lodewijks, G., 2015. Service network design for an intermodal container network with flexible transit times and the possibility of using subcontracted transport. International Journal of Shipping and Transport Logistics 7, 457–478.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics 20, 53–65.

SteadieSeifi, M., Dellaert, N.P., Nuijten, W., Van Woensel, T., Raoufi, R., 2014. Multimodal freight transportation planning: A literature review. European journal of operational research 233, 1–15.

Sun, Y., Zhang, G., Hong, Z., Dong, K., 2018. How uncertain information on service capacity influences the intermodal routing decision: A fuzzy programming perspective. Information 9, 24.

Wang, X., Crainic, T.G., Wallace, S.W., 2019. Stochastic network design for planning scheduled transportation services: The value of deterministic solutions. INFORMS Journal on Computing 31, 153–170.

Wang, Z., Qi, M., 2020. Robust service network design under demand uncertainty. Transportation Science 54, 676–689.

Ypsilantis, P., Zuidwijk, R., 2019. Collaborative fleet deployment and routing for sustainable transport. Sustainability 11, 5666.

Zhang, R., Huang, C., Feng, X., 2020. Empty container repositioning with foldable containers in a river transport network considering the limitations of bridge heights. Transportation Research Part A: Policy and Practice 133, 197–213.

Zheng, Y., Kim, A.M., 2017. Rethinking business-as-usual: Mackenzie river freight transport in the context of climate change impacts in northern canada. Transportation Research Part D: Transport and Environment 53, 276–289.

Appendices

A.1 The Deterministic WL-SSND-RRM model

In the following, we present the deterministic formulation of the WL-SSND-RRM problem. This model assumes constant (though not necessarily identical) water levels over the planning horizon for each physical link in the transportation network, based on point forecasts. A point forecast provides a single predicted value for the water level, representing situations with low variability in long-term water level predictions. Given these predicted water levels, the model aims to maximize the expected carrier's revenue while designing the service network and determining the optimal routing of demand, ensuring demand fulfillment, and addressing the needs of both regular and selected spot shippers. In this modeling framework, decision variables are as follows:

- **Service Selection**
  - $y(\sigma)$**:** Determines which transportation services to select.
- **Shipper Selection**
  - $\xi(k)$**:** Determines the percentage of the volume of partial-spot shipper demand $k \in \mathcal{K}^{\mathrm{P}}$ that is selected and will be serviced.
  - $\zeta(k)$**:** Determines whether full-spot shipper demand $k \in \mathcal{K}^{\mathrm{F}}$ will be serviced or not.
- **Resource Management**
  - $z(v, a_{\eta_t})$**:** Determines the number of temporarily idle vessels of type v waiting at holding arc a_{η_t} for the departure of the next service they support.
  - $B(v)$**:** Determines the total number of vessels of type v used in the service plan. Due to the circular nature of the schedule, $B(v)$ remains constant across all time periods, although at any given period, vessels may either be moving or idle in ports.
- **Demand Flow Distribution**
  - $x(k, a_{l_i}(\sigma))$**:** Volume of demand $k \in \mathcal{K}$ transported on the i^{th} leg of service σ.
  - $\mathrm{x}^{\mathrm{out}}(k, a_{l_i}(\sigma))$**:** Volume of demand $k \in \mathcal{K}$ unloaded from the i^{th} leg of service σ.
  - $\mathrm{x}^{\mathrm{in}}(k, a_{l_i}(\sigma))$**:** Volume of demand $k \in \mathcal{K}$ loaded onto the i^{th} leg of service σ.
  - $\mathrm{x}^{\mathrm{hold}}(k, a_{\eta_t})$**:** Volume of demand $k \in \mathcal{K}$ held at terminal η during the time interval $(t, t+1)$.

The determinstic formulation of the WL-SSND-RRM problem can then be written as:

$$\begin{aligned}\max \sum_{k\in\mathcal{K}^{\mathrm{R}}} \phi(k)d(k) + \sum_{k\in\mathcal{K}^{\mathrm{P}}} \phi(k)\xi(k)d(k) + \sum_{k\in\mathcal{K}^{\mathrm{F}}} \phi(k)\zeta(k)d(k) \\ - \sum_{v\in\mathcal{V}} \mu(v)(F_v - B(v)) - \sum_{\sigma\in\Sigma} f(\sigma)y(\sigma) - \sum_{a_{\eta_t}\in\mathcal{A}^H} \sum_{v\in\mathcal{V}} h(\eta,v)z(v,a_{\eta_t}) \\ - \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M} \sum_{k\in\mathcal{K}} c_i(\gamma(k),v(\sigma))x(k,a_{l_i}(\sigma)) - \sum_{a_{\eta_t}\in\mathcal{A}^H} \sum_{k\in\mathcal{K}} c(\eta,\gamma(k))\mathrm{x}^{\mathrm{hold}}(k,a_{\eta_t}) \\ - \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M} \sum_{k\in\mathcal{K}} \kappa(\eta_i(\sigma),\gamma(k))(\mathrm{x}^{\mathrm{in}}(k,a_{l_i}(\sigma)) + \mathrm{x}^{\mathrm{out}}(k,a_{l_i}(\sigma)))\end{aligned} \tag{1}$$

Subject to

$$\mathrm{x}^{\mathrm{hold}}(k,a_{\eta_t}) + \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M,(\eta_{i-1}=\eta,\alpha(\eta_{i-1})=t)} \mathrm{x}^{\mathrm{in}}(k,a_{l_i}(\sigma)) = \begin{cases} d(k), & \forall k\in\mathcal{K}^{\mathrm{R}},\eta=O(k),t=\alpha(k) \\ \xi(k)d(k), & \forall k\in\mathcal{K}^{\mathrm{P}},\eta=O(k),t=\alpha(k) \\ \zeta(k)d(k), & \forall k\in\mathcal{K}^{\mathrm{F}},\eta=O(k),t=\alpha(k) \end{cases} \tag{2}$$

$$\sum_{\alpha(k)<t\le\beta(k)} \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M,(\eta_i=\eta,\beta(\eta_i)=t)} \mathrm{x}^{\mathrm{out}}(k,a_{l_i}(\sigma)) = \begin{cases} d(k), & \forall k\in\mathcal{K}^{\mathrm{R}},\eta=D(k), \\ \xi(k)d(k), & \forall k\in\mathcal{K}^{\mathrm{P}},\eta=D(k) \\ \zeta(k)d(k), & \forall k\in\mathcal{K}^{\mathrm{F}},\eta=D(k) \end{cases} \tag{3}$$

$$\begin{aligned}\mathrm{x}^{\mathrm{hold}}(k,a_{\eta_{t-1}}) + \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M(\eta_i=\eta,\beta(\eta_i)=t)} \mathrm{x}^{\mathrm{out}}(k,a_{l_i}(\sigma)) \\ -\mathrm{x}^{\mathrm{hold}}(k,a_{\eta_t}) - \sum_{a_{l_i}(\sigma)\in\mathcal{A}^M((\eta_{i-1}=\eta,\alpha(\eta_{i-1})=t)} \mathrm{x}^{\mathrm{in}}(k,a_{l_i}(\sigma)) = 0 \\ \forall(\eta,t)\neq(O(k),\alpha(k))\forall\eta\neq D(k),\forall k\in\mathcal{K}\end{aligned} \tag{4}$$

$$\mathrm{x}^{\mathrm{in}}(k,a_{l_i}(\sigma)) - x(k,a_{l_i}(\sigma)) = 0, \forall a_{l_i}(\sigma)\in\mathcal{A}^M,\ \eta_{i-1}(\sigma)=O(\sigma),\ k\in\mathcal{K} \tag{5}$$

$$x(k,a_{l_i}(\sigma)) - \mathrm{x}^{\mathrm{out}}(k,a_{l_i}(\sigma)) = 0, \forall a_{l_i}(\sigma)\in\mathcal{A}^M,\ \eta_i(\sigma)=D(\sigma),\ k\in\mathcal{K} \tag{6}$$

$$\begin{aligned}x(k,a_{l_{i-1}}(\sigma)) - \mathrm{x}^{\mathrm{out}}(k,a_{l_{i-1}}(\sigma)) + \mathrm{x}^{\mathrm{in}}(k,a_{l_i}(\sigma)) - x(k,a_{l_i}(\sigma)) = 0, \\ \forall\sigma\in\Sigma,\ \eta_{i-1}\neq O(\sigma),\eta_i\neq D(\sigma),\ k\in\mathcal{K}\end{aligned} \tag{7}$$

$$\omega\sum_{k\in\mathcal{K}} x(k,a_{l_i}(\sigma)) \le cap_w(v(\sigma))y(\sigma), \quad \forall\sigma\in\Sigma, a_{l_i}(\sigma)\in\mathcal{A}^M \tag{8}$$

$$\sum_{k\in\mathcal{K}} x(k,a_{l_i}(\sigma)) \leq cap_{vol}(v(\sigma))y(\sigma), \quad \forall\sigma\in\Sigma, a_{l_i}(\sigma)\in\mathcal{A}^M \tag{9}$$

$$\theta(v(\sigma))\omega\sum_{k\in\mathcal{K}} x(k,a_{l_i}(\sigma)) + dh^-(v(\sigma)) \leq p_{a_{l_i}(\sigma)} \quad \forall\sigma\in\Sigma, a_{l_i}(\sigma)\in\mathcal{A}^M \tag{10}$$

$$B(v) = \sum_{\eta\in\mathcal{N}^{\mathrm{ph}}} z(v,a_{\eta_0}) + \sum_{\sigma\in\Lambda_{0l}} y(\sigma), \quad \forall v\in\mathcal{V} \tag{11}$$

$$B(v) \leq F_v, \quad \forall v\in\mathcal{V} \tag{12}$$

$$\sum_{\sigma\in\Sigma^-_{\eta tv}} y(\sigma) + z(v,a_{\eta_{t-1}}) = \sum_{\sigma\in\Sigma^+_{\eta tv}} y(\sigma) + z(v,a_{\eta_t}) \quad \forall v\in\mathcal{V},\ a_{\eta_{t-1}}, a_{\eta_t}\in\mathcal{A}^H \tag{13}$$

$$\sum_{v\in\mathcal{V}} len(v)(\sum_{\sigma\in\Sigma^-_{\eta tv}} y_\sigma + z(v,a_{\eta_{t-1}})) \leq Q_{a_{\eta t}}, \quad \forall a_{\eta t}\in\mathcal{A}^H \tag{14}$$

$$y(\sigma)\in\{0,1\} \quad \forall\sigma\in\Sigma \tag{15}$$
$$\xi(k)\in[0,1] \quad \forall k\in\mathcal{K}^{\mathrm{P}} \tag{16}$$
$$\zeta(k)\in\{0,1\}\forall k\in\mathcal{K}^{\mathrm{F}} \tag{17}$$
$$z(v,a_{\eta_t})\geq 0 \quad \forall v\in\mathcal{V}, \quad a_{\eta_t}\in\mathcal{A}^H \tag{18}$$
$$B(v)\geq 0, \text{integer} \quad \forall v\in\mathcal{V} \tag{19}$$
$$x(k,a_{l_i}(\sigma))\geq 0 \quad \forall k\in\mathcal{K},\ a_{l_i}(\sigma)\in\mathcal{A}^M \tag{20}$$
$$\mathrm{x}^{\mathrm{out}}(k,a_{l_i}(\sigma))\geq 0 \quad \forall k\in\mathcal{K},\ a_{l_i}(\sigma)\in\mathcal{A}^M \tag{21}$$
$$\mathrm{x}^{\mathrm{in}}(k,a_{l_i}(\sigma))\geq 0 \quad \forall k\in\mathcal{K},\ a_{l_i}(\sigma)\in\mathcal{A}^M \tag{22}$$
$$\mathrm{x}^{\mathrm{hold}}(k,a_{\eta_t})\geq 0 \quad \forall k\in\mathcal{K},\ a_{\eta_t}\in\mathcal{A}^H. \tag{23}$$

The objective function (1) maximizes the net profit of a carrier, where the first three terms correspond to the revenue obtained by servicing the complete demand of regular shippers, the selected proportion of demand of the partial-spot shippers, and the complete demand of the selected full-spot shippers respectively. Remark that the first term (revenue obtained by servicing the complete demand of regular shippers) is kept in the formulation to have the objective function as a homogeneous mathematical expression. The following terms stand for the activity and time-related costs of operating the selected service network and resource routes, that is, the penalty cost of having but not using vessels (never assigned to service during the entire schedule length), the fixed cost of setting up and operating services, the cost of the vessels

idling at a port waiting for their next service departure, the cost of transporting containers on services, and the cost of holding and handling containers in terminals.

Equations (2), (3), and (4) are flow-conservation constraints for containers of all shipper types, at their particular origins, destinations, and intermediary nodes, respectively. Similarly, Equations (5), (6) and (7) enforce the conservation of container flows, for all shipper types, on each service at its origin, destination and intermediary stops, respectively. (8) and (9) guarantee that the weight and volume of demand $k \in \mathcal{K}$ transported by service σ on its leg i do not exceed the nominal capacity of vessel v performing service σ. Constraint (10) assures that the draught of vessels executing i^{th}leg of service σ is less than or equal to water level on that leg. Equation (11) computes the number of vessels used in the plan as the sum of vessels idling in ports or moving between them performing services. Due to the resource management concerns and the resulting circular vessel routes, $B(v)$ is the same at all periods, only the relative proportion of idle versus active vessels being different at different time periods. We therefore compute this number for the first period, i.e., $t = 0$, the set $\Lambda_{0l} = \{\sigma \in \Sigma, v(\sigma) = v | (\alpha_{n(\sigma)} \mod T) < \beta_0(\sigma) \text{ and } \beta_0(\sigma) \geq 0\} \subseteq \Sigma$ containing all services, of the appropriate vessel type, that operate one of its legs during the first period. Constraints (12) enforce the fleet size for each vessel type, while Equations (13) are the so-called design-balance constraints, enforcing the vehicle-flow conservation at terminals (the number of services and vessels entering a node equals the number exiting the node), where sets $\Sigma^-_{\eta tv}$ and $\Sigma^+_{\eta tv}$

$$\Sigma^-_{\eta tv} = \{\sigma \in \sigma \mid D(\sigma) = \eta, \beta_n(\sigma) = t, v(\sigma) = v\} \quad (24)$$

$$\Sigma^+_{\eta tv} = \{\sigma \in \Sigma \mid O(\sigma) = \eta, \alpha_0(\sigma) = t, v(\sigma) = v\} \quad (25)$$

group the services with a vessel type v that arrive at their destination or depart from their origin, respectively. Finally, Constraints (14) enforce the terminal berthing capacity at each time period. Decision-variable domains are defined by Constraints (15) - (23).

A.2 Test Instances

For our study, the schedule length is defined as a one-week period (7 days), divided into 14 equal half-day time periods. This division reflects the typical operational schedules observed in inland waterway transportation, where service-related activities, such as departures and arrivals, are commonly organized into morning and afternoon slots during workdays.

This section provides an overview of the test instances used in our study. The first subsection details the physical network topology and its relevant features related to terminals and waterways. The second subsection outlines the potential service network generation process, and the third subsection describes the generation of demand data.

Physical network and its relevant features. We generate a set of test instances based on a linear physical network topology with four sequentially connected terminals (see Figure 5). This linear network represents a common network type, observed in Canadian and European waterway systems, characterized by the sequential connectivity of major ports along their respective waterways. For instance, the most important ports along the Saint Lawrence River—Montréal, Sept-Îles, Québec City, and Port-Cartier—are linked in sequence, facilitating

the continuous flow of goods along this vital corridor. Similarly, the Rhine River, divided into the Upper, Middle, and Lower Rhine, forms a linear and sequential network of interconnected ports that support European trade. In the Upper Rhine, key ports such as Basel, Strasbourg, Mannheim, and Karlsruhe are linked in a direct sequence, ensuring the smooth flow of goods along this section. The Middle Rhine connects ports like Koblenz, Mainz, Bingen, and Wiesbaden, while the Lower Rhine features ports such as Duisburg, Düsseldorf, Wesel, and Emmerich, all maintaining their linear connectivity to support regional and international trade.

Terminal berthing capacities are generated, ranging from 700 to 1000 meters. Considering vessel lengths ranging from 100 to 200 meters, this capacity allows for the simultaneous berthing of multiple vessels. This capability is supported by the port's ability to handle loading and unloading operations and the presence of extensive parallel tracks and intermodal facilities. Costs associated with terminal operations are considered as follows: the loading and unloading of containers incur a cost of 2 units per operation, terminal activities such as storage are charged to carriers at a rate of 3 units per container per period, and vessel holding costs at terminals are set at 5 units per period.

Additionally, for this physical network, we account for stochastic water levels at critical points along the waterways. Scenarios capturing this stochastic behavior are generated using the copula-based method proposed by (Kaut, 2014), ensuring that the statistical properties of the water levels adhere to the target marginal distributions and the correlation matrix between terminal pairs.

The marginal distributions and the correlation matrix are derived from historical data that capture the random variability of water levels across the corridor waterways during the season under consideration for planning. To model the marginal distributions, we focus on the range of water levels defined by (Christodoulou et al., 2020), spanning from 150 cm to 350 cm. Among the candidate distributions evaluated, the Beta distribution is identified as the most suitable. These distributions allow for capturing different types of skewness in the random variations of water level values. In our case, during periods of drought, one may observe a right-skewed distribution, where the density of water level values is higher below the mean (250 cm), indicating that low water level values occur more frequently. Meanwhile, higher water level values than the mean occur less frequently and have lower density, which is represented by the longer right tail capturing rare high-water occurrences. Specifically, the Beta(2,5) distribution is employed due to its ability to represent right-skewed distributions. This distribution captures the tendency for most values to be below the mean while still allowing for infrequent but high water levels.

The correlation matrix is generated based on the physical distances between terminals, with correlations decreasing as the distance between terminal pairs increases. In the linear network, the correlation matrix reflects the proximity of nodes. As depicted in Fig. 5, for example, Node 0 shows the strongest correlation with Node 1 (0.75), followed by Node 2 (0.5) and Node 3 (0.25). This pattern reflects the diminishing influence of water levels as the distance between nodes increases.

The copula-based heuristic uses the mean and standard deviation of the Beta distribution, along with the defined correlation matrix, to generate a predefined number of scenarios $|\mathcal{S}|$, each with equal probability $\left(\rho^s = \frac{1}{|\mathcal{S}|}\right)$.

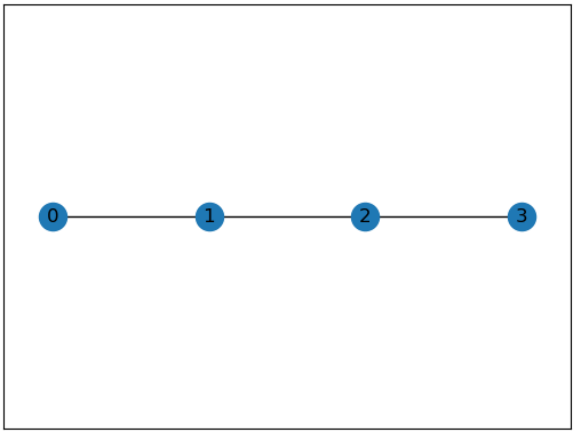


Figure 5: Physical network

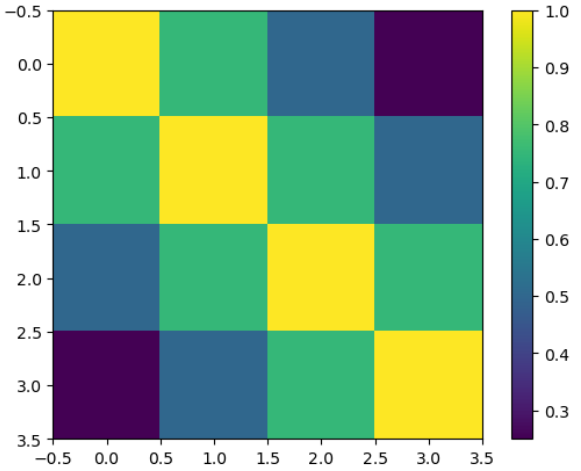


Figure 6: Correlation matrix

Service network generation. Potential services are tailored to the linear network topology, considering all possible non-stop and one-stop routes. The services are scheduled to start at various times within the weekly cycle, which is divided into 14 periods as indicated previously. Out of these 14 periods, 9 distinct start times are selected.

The total number of possible non-stop and one-stop routes is calculated as follows:

For non-stop routes: Two terminals are selected as the origin and destination terminals from the 4 existing terminals. This results in a total of: $\binom{4}{2} = 6$ non-stop routes.

For one-stop routes: Three terminals are selected: one as the origin, one as the destination, and one as the intermediate terminal, from the 4 existing terminals. This results in a total of:$\binom{4}{3} = 4$ one-stop routes.

Since both directions are considered, the total number of non-stop and one-stop routes becomes doubled, resulting in a total of 20 routes.

Considering these 20 routes, the total number of services is calculated by multiplying the number of routes by the number of start times, resulting in 180 services. It is assumed that each of the 180 considered services can be operated using either a large or a small vessel. Thus, we define a total of 360 services, where each service is further characterized by the type of vessel assigned to perform it. The duration of each service is calculated based on the travel time between terminals and the time spent at each origin, destination, and intermediate terminal, which is assumed to be one period per stop.

Travel times between terminals are defined proportionally to the distance between them, reflecting the linear network topology. Adjacent terminals have the shortest travel times, and as the distance increases, the travel times increase proportionally. Large vessels are faster, with travel times that are half those of small vessels. For instance, if a large vessel takes 1 unit of time to travel between two adjacent terminals, a small vessel requires 2 units for the same route. This difference ensures variability in operations and highlights the impact of vessel type on service durations.

Dimensions of each vessel type are as follows: Large Vessels: Measuring 200 meters in length, with a minimum draught of 135 centimeters, these vessels support a maximum weight of 500 tons and can carry up to 50 TEUs. The draught-weight factor for large vessels is 0.4. Small Vessels: Half the length of large vessels (100 meters), with a minimum draught of 100 centimeters, these vessels can carry up to 200 tons and 20 TEUs. The draught-weight factor for small vessels is 0.3. Operating a large vessel incurs approximately twice the cost of operating a small one, reflecting economies of scale.

It is noteworthy that both vessel types are capable of navigating the entire network in both directions when they are empty because the minimum water level is 150 cm.

Demand Generation. We generate five distinct demand instances for the experimental results, each consisting of 55 to 66 demands, resulting in a total of 298 demands. This approach minimizes potential biases or anomalies from individual instances, captures a wider range of demand conditions, and provides more generalizable and reliable insights into the system's behavior under different water level scenarios.

Demand is for the linear network based on randomly selected origin-destination pairs, ensuring a broad coverage of possible transportation requests. The demands are characterized by their volume randomly generated between 5 and 25 units.

Recall that the demand of regular, partial spot, and full spot shippers is classified into two types based on delivery time requirements (i.e., express or standard), the unit revenue being twice higher for express deliveries.

The desired time window for picking up each request from its origin is a value uniformly generated between one and seven. The duration for delivering the request to its destination depends on the type of the request. For an express request, the delivery duration is generated randomly between one and three. The delivery duration of a standard request is twice as large as the duration for an urgent request with the same origin and destination.

We complete this section by introducing the instance categories labeled Instance-1 through Instance-7. Each category includes three demand instances, with two recourse costs, denoted by $b(k)$ and $c_i(\gamma(k), v(\sigma))$, randomly selected within specified ranges. For the 2-SPSDM model, $b(k)$ is randomly chosen between 5 and 60 units, representing penalty costs for failing to meet demands. These penalty costs are set to be higher than other costs, as they have distinct characteristics, such as fines. For the 2-SPDFA model, $c_i(\gamma(k), v(\sigma))$ represents the transportation cost of a container for one leg of service, randomly selected between 3 and 7 units.

A.2.1 Stability test

While increasing the number of scenarios improves the representation of the water level distribution, it also complicates the process of obtaining an optimal solution. Our scenario generation method is inherently random, meaning that if the procedure is rerun with the same inputs, the resulting scenario tree varies. To ensure robustness, we use the approach proposed by ((Kaut et al., 2007)). We solve our two-stage stochastic models with a specific number of scenarios across 10 different randomly generated scenario trees with Gurobi. We then calculate the average and standard deviation of the objective function values. By repeating this process for

different scenario counts, we study the effect of the number of scenarios on the stability of the results.

Table 5: Stability Results for 2-SPSDM model

Scenario Size	**Overall Average**	**Overall SD**
10	178943.86	4788.49
20	174989.42	5576.04
30	172455.44	3950.13
40	170642.14	3138.39

Table 6: Stability Results for 2-SPDFA model

Scenario Size	**Overall Average**	**Overall SD**
10	113088.30	2315.56
20	110212.41	1651.58
30	109274.73	1311.68
40	108167.22	1308.50

Table 5 presents the optimal results obtained from the 2-SPSDM model, indicating that as the number of scenarios increases, the Standard Deviation (SD) decreases, which translates to an increase in in-sample stability. As the number of scenarios rises, the Coefficient of Variation (CV)—the ratio of the standard deviation to the average—drops to 2.29%. However, when the number of scenarios increases from 30 to 40, the CV further decreases to 1.83%. Considering the computational cost of using 40 scenarios compared to 30, alongside the slight reduction in the CV value, we determine that 30 is the appropriate choice for the following experiments.

In Table 6, the best-known solutions for the 2-SPDFA model are presented. We set a runtime limit of 6 hours and obtain these results. The results indicate a similar trend to the first stochastic model. Based on these findings, we identify 20 scenarios as the best option for this model, achieving a CV of 1.49%.