CIRRELT

CIRRELT-2026-11

The Technician Routing and Scheduling Problem with Skills and Time-Sensitive Returns under Uncertainty

Milad Elyasi
Amira Dems
Yossiri Adulyasak
Okan Arslan
Jean-François Cordeau

March 2026

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

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

The Technician Routing and Scheduling Problem with Skills and Time-Sensitive Returns under Uncertainty

Milad Elyasi[1,2*], Amira Dems[3], Yossiri Adulyasak[2], Okan Arslan[2], Jean-François Cordeau[2]

[1] Industrial Engineering Department, Özyeğin University, Istanbul, Turkey

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and HEC Montréal, Montreal (QC)

[3] Institut de recherche d'Hydro-Québec, Varennes (QC)

Abstract. We study the scheduling and routing of technician teams with the objective of maximizing the overall benefit of services provided while satisfying operational constraints on skills, workloads, routing, and working hours. The problem is motivated by a real-world collaboration with Hydro-Québec, where large-scale field operations require daily technician dispatching under tight time, skill, and uncertainty considerations. The problem is formulated over a finite multi-period planning horizon, differentiating between prioritized and optional customer visits and incorporating diminishing service benefits over time. Technician skill heterogeneity, vehicle capacity limits, and travel and service times are explicitly modeled, with stochastic extensions capturing uncertainty through chance constraints. As the main methodological contribution, we develop a tailored Logic-Based Benders Decomposition (LBBD) algorithm that decomposes the problem into an assignment-based master problem and routing feasibility subproblems. Routing feasibility is verified via dedicated TSP solvers in deterministic settings and conic-quadratic formulations under uncertainty, enabling scalability without sacrificing solution quality. Extensive computational experiments on 210 benchmark instances demonstrate that the proposed LBBD substantially outperforms a Branch-and-Cut benchmark. LBBD solves 202 instances to optimality, compared to 161 for Branch-and-Cut, achieves significantly lower average optimality gaps (6.8% versus 31.8%), and reduced average computation time (1616 s versus 2493 s). A real-world case study with 200 customers over a multi-period horizon confirms the practical applicability of the approach. The results show that LBBD produces robust and operationally viable schedules under uncertainty while maintaining high service coverage and improved workforce utilization with moderate computational effort.

Keywords: stochastic programming, technician routing, technician scheduling

---

* Corresponding author: milad.elyasi@ozyegin.edu.tr

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

1 Introduction

Industries such as telecommunications, utilities, and healthcare depend on mobile workforces to perform installations, maintenance, and repairs at customer locations (Bughin et al., 2013). In recent years, service companies that rely on field technicians have faced increasing pressure to enhance operational efficiency due to rising expectations for prompt and reliable service. Meeting these expectations requires effective routing and scheduling of technicians whose skills and availability directly affect service quality. The resulting Technician Routing and Scheduling Problem (TRSP) focuses on optimizing technician assignments to customer locations while accounting for skills, availability, and travel time.

Early research on TRSP primarily examined simplified scenarios in which technicians share identical skills and customers are treated uniformly (Pillac et al., 2013). Such models overlook key real-world factors including skill mismatches, customer prioritization driven by contractual or profit considerations, and operational constraints such as maximum work hours or team requirements. In practice, many service operations require technicians to work in teams where task feasibility depends on the crew composition. Importantly, not all requirements behave like pooled capacities: some competencies can be combined across technicians (e.g., aggregate proficiency or manpower-based capability), while others correspond to individual qualifications that must be held by at least one team member (e.g., certifications, permits, or safety clearances). Capturing this hybrid nature of skills adds substantial complexity, as schedules must account not only for individual competencies and feasible team compositions, but also for the correct interpretation of skill feasibility. As companies strive to improve operational efficiency, more sophisticated models capable of handling technician heterogeneity, customer prioritization, and regulatory constraints have become essential. Efficient routing reduces travel times, enabling companies to serve more customers. Furthermore, optimized scheduling improves customer satisfaction by ensuring that services are delivered, which provides a competitive advantage in markets with high customer expectations. Additionally, it enhances service quality by ensuring that technicians with the required skills are assigned tasks, reducing service time and ensuring that customer needs are met more effectively (Li et al., 2020). However, despite these benefits, the complexities involved in scheduling and routing technicians—such as heterogeneous skills, customer prioritization, and operational constraints—present significant challenges that must be addressed to achieve these efficiencies.

Customers can be categorized into prioritized customers who must be serviced within the planning horizon and optional ones who can be serviced if resources allow, adding further complexity to scheduling decisions (Chen et al., 2016b). To address these challenges, we introduce the Technician Routing and Scheduling Problem with Skills and Time-Sensitive Returns under Uncertainty (TRSP-STSU). We present a mathematical model that incorporates team-based staffing with a hybrid skill-feasibility structure, time-sensitive service benefits, and uncertainty in travel and service times, and we develop a Logic-Based Benders Decomposition (LBBD) solution algorithm. The LBBD method decomposes the problem into a master scheduling component and routing subproblems, enabling efficient handling of team assignments, skill diversity, and customer prioritization under uncertainty.

To provide a benchmark for evaluating the effectiveness of the proposed LBBD approach, we also implement an additional exact solution method based on a Branch-and-Cut (BC) framework. The BC algorithm solves the integrated formulation directly by embedding subtour elimination constraints and other problem-specific cuts within a branch-and-bound process, progressively tightening the relaxation as the search tree is explored. While BC is capable of enforcing all operational constraints in a unified manner and guarantees optimality, its performance typically deteriorates as problem size and routing complexity increase. This comparison allows us to assess the scalability, robustness, and computational advantages of the LBBD algorithm relative to a strong monolithic exact method.

The remainder of the paper is organized as follows. Section 2 provides the literature review. Section 3 formally defines the problem and presents the mathematical formulations. Section 4 describes the decomposition-based algorithms. Section 5 presents computational results, compares our approach with existing methods, and discusses practical implications. Section 6 concludes the paper and outlines directions for future research.

2 Literature Review

In the field of TRSP, extensive research has addressed the challenges of efficiently allocating resources while meeting operational constraints (Mathlouthi et al., 2021). The TRSP originates from the broader class of vehicle routing problems and has evolved in response to increasing demand for improved service delivery, reduced operational costs, and enhanced customer satisfaction in service-intensive industries (Pourjavad and Almehdawe, 2022). For a comprehensive review of TRSP and related workforce routing problems, we refer the readers to the survey by Cabrera et al. (2026) that synthesizes recent advances in modeling approaches, solution methodologies, and application domains. The TRSP represents a specialized workforce scheduling and routing problem (Castillo-Salazar et al., 2016), characterized by heterogeneous technician skill sets and diverse service tasks. An important feature of these problems is that service times at customer sites typically dominate travel times. Early work treated the challenge as a scheduling problem without routing; for example, Haugen and Hill (1999) proposed a Mixed-Integer Linear Programming (MILP) formulation and simulation model to enhance service quality through improved technician scheduling. Complexity increases substantially when routing is incorporated, as shown by Kovacs et al. (2012). Likewise, Pillac et al. (2013) examine a TRSP involving time windows, required skills, tools, and spare parts, and propose a matheuristic combining a constructive heuristic, a parallel Adaptive Large Neighborhood Search (ALNS), and a mathematical programming–based post-optimization procedure.

Incorporating a multi-period planning horizon into the TRSP enables more effective scheduling and routing over extended timeframes, improving resource allocation and task completion (Nowak and Szufel, 2024). Bostel et al. (2008) formulate a multi-period TRSP for a water distribution company using a rolling-horizon schedule updated daily, and optimize the initial plan with a memetic algorithm combined with a column generation and branch-and-bound heuristic. Zamorano and Stolletz (2017) study a TRSP for an external maintenance provider where technicians with different skills form teams to complete skill-constrained tasks with multi-day time windows; they exploit problem structure to develop alternative formulations for a branch-and-price algorithm. Similarly, Chen et al. (2018) address a variant where only current-day tasks are known and future tasks remain uncertain, proposing an approximate dynamic programming approach. Khattara et al. (2017) introduce a variable neighborhood search that outperforms earlier memetic algorithms, particularly for large-scale telecommunication maintenance operations. For dynamic, real-time reoptimization, Pillac et al. (2018) develop a fast parallel ALNS that efficiently updates schedules and routes. Additional examples of multi-period TRSPs can be found in Barrera et al. (2012) and in Chen et al. (2016b).

Skill levels and task requirements are central to many TRSP variants (Pekel, 2020). Skill diversity plays a decisive role in task assignment and routing. Early work by Taillard et al. (1997) formally incorporates skill considerations into routing, establishing the foundation for competency-based technician assignment. Since then, numerous studies emphasize the importance of matching technician competencies with task demands, as variations in skill levels directly influence route structure, task sequencing, and overall service feasibility (Moradi, 2020). Managing diverse skill sets becomes particularly challenging in large-scale settings, where balancing skill utilization and task allocation is essential for operational efficiency.

The TRSP is also closely linked to routing problems with time windows (Moradi, 2020; Paradiso et al., 2020), which add an additional layer of complexity. Mathlouthi et al. (2021) study a TRSP for electronic transactions equipment maintenance involving multiple time windows and special parts requirements, proposing a tabu search approach with adaptive memory. Such studies underscore the importance of managing heterogeneous tasks with specific operational requirements. Kovacs et al. (2012) address a TRSP where technicians with varying skill levels are assigned to diverse tasks, aiming to minimize routing and outsourcing costs through an adaptive large neighborhood search. More recently, Chen et al. (2024) model a multi-period dynamic TRSP as a Markov decision process, explicitly capturing individualized technician learning and the heterogeneity of tasks and skills. In contrast to this stream of work, we abstract away from explicit service time windows and instead model temporal urgency through diminishing returns over the planning horizon, while enforcing feasibility via per-period working-time limits.

Uncertainty is inherent in real-world TRSP applications and significantly affects operational performance. Variability in travel times—arising from factors such as traffic or weather—and fluctuations in service durations due to task complexity can compromise both efficiency and schedule feasibility. Addressing such uncertainties requires robust or stochastic optimization approaches capable of adapting to real-time deviations. Chen et al. (2016a) propose a stochastic TRSP that incorporates dynamic variability in travel and service times, enabling more realistic planning. Likewise, Zamorano and Stolletz (2017) develop a branch-and-cut algorithm that accounts for stochastic task durations, ensuring flexibility when actual execution times deviate from expected values. Incorporating these uncertainties enhances the adaptability of TRSP solutions and improves service performance in environments where perfect information is rarely available.

Despite extensive research on various TRSP variants, several critical gaps remain, particularly with respect to the complexities of real-world applications. Many existing studies rely on simplifying assumptions—such as homogeneous technician skills or uniform treatment of customers—which overlook skill mismatches and the need to prioritize customers based on contractual obligations or economic value. Moreover, important operational constraints, including maximum working hours, team composition requirements, and diminishing returns over time that reflect service urgency, are often omitted or only partially modeled. Uncertainty in travel and service times, arising from traffic variability or task complexity, further complicates scheduling decisions and is rarely captured in a unified and tractable optimization framework.

To address these challenges, we introduce the technician routing and scheduling problem with skill diversity and diminishing benefits under uncertainty (TRSP-STSU). The proposed model integrates heterogeneous technician skills, customer prioritization, a multi-period planning horizon, and operational uncertainty within a single formulation. To effectively solve large-scale instances of this problem, we develop decomposition-based solution approaches that explicitly exploit its structural properties. In particular, we design a tailored LBBD algorithm that separates assignment and scheduling decisions from routing feasibility, enabling efficient handling of team-based assignments, skill constraints, and time-dependent benefits. The LBBD framework is strengthened through several algorithmic enhancements, including uniform preprocessing rules that eliminate provably infeasible assignments, no-good cuts that prevent recurring infeasible patterns, sequence-independent duration bounds that rule out time-infeasible visit combinations, and valid inequalities that reduce symmetry and tighten the master problem relaxation. These enhancements significantly reduce the search space and improve convergence without restricting feasible solutions. To assess the effectiveness and robustness of the proposed LBBD algorithm, we benchmark it against an alternative exact decomposition-based approach, namely a BC algorithm, through extensive computational experiments.

The contributions of this study are threefold: (i) we propose a mathematical formulation for the TRSP-STSU that captures team assignments, skill diversity, customer prioritization, diminishing returns over time, and key operational constraints under uncertainty; (ii) we develop and enhance a tailored LBBD algorithm, incorporating preprocessing, specialized Benders cuts, and routing-aware valid inequalities, and systematically benchmark its performance against a strong BC baseline; and (iii) we analyze the impact of skill diversity, customer prioritization, uncertainty levels, and labor policies on operational efficiency and solution quality, providing quantitative insights into the trade-offs inherent in technician routing and scheduling decisions.

3 Problem Definition

In this section, we present the problem of scheduling and routing teams of technicians to maximize overall benefit while satisfying various operational constraints. The primary objective is to effectively allocate resources to service a set of customers within a given planning horizon, which is divided into discrete time periods. Technicians must complete their assigned tasks within these time periods, adhering to work hour regulations and ensuring efficient use of resources.

Each team must collectively possess the required skills to service an assigned customer. In practice, skill requirements may be heterogeneous in nature: some requirements behave like *divisible capacities* that can be pooled across the crew (e.g., aggregate proficiency or manpower-based capability), while others behave like *qualifications* that must be held by at least one technician (e.g., certifications, permits, or safety clearances). We therefore adopt a *hybrid* skill-feasibility structure by partitioning the set of skill types into additive skills $\mathcal{M}^{\text{add}}$ and qualification (coverage) skills $\mathcal{M}^{\text{cov}}$, with $\mathcal{M} = \mathcal{M}^{\text{add}} \cup \mathcal{M}^{\text{cov}}$ and $\mathcal{M}^{\text{add}} \cap \mathcal{M}^{\text{cov}} = \emptyset$. For additive skills $m \in \mathcal{M}^{\text{add}}$, feasibility is checked by aggregating the skill levels R_{lm} of the technicians assigned to a vehicle, i.e., the team as a whole must meet or exceed the minimum required level R'_{im}. For qualification skills $m \in \mathcal{M}^{\text{cov}}$, feasibility requires that at least one technician on the team individually satisfies the requirement (e.g., $R_{lm} \geq R'_{im}$).

The time required to travel between customer locations and the time spent servicing each customer are significant factors affecting scheduling and routing decisions. These times contribute to the total working time of each technician team, which must not exceed the maximum allowable work hours per period (T^{max}) in accordance with labor regulations and to prevent overwork. The total working time therefore includes both travel time between successive customers and the service time spent on-site. Customers are categorized into two sets: prioritized customers, denoted by $\mathcal{I}^p$, and optional customers, denoted by $\mathcal{I}^o$, with $\mathcal{I} = \mathcal{I}^o \cup \mathcal{I}^p$. Each prioritized customer must be serviced exactly once within the planning horizon, whereas optional customers may be serviced if sufficient time and resources remain, potentially increasing overall benefit. Optional customers who are not visited may be elevated to prioritized status in subsequent planning horizons. A further practical consideration is the *urgency of service requests*, reflecting the fact that responding promptly to pressing issues yields higher operational and customer value. Service delays can diminish customer satisfaction, increase the likelihood of revenue losses, and generate additional costs. Incorporating urgency explicitly allows the model to prioritize tasks where early intervention is most beneficial, thereby maximizing customer satisfaction and operational efficiency.

To mathematically capture the diminishing returns associated with delayed service, we define the parameter G_{it} as the benefit of serving customer $i \in \mathcal{I}$ in period $t \in \mathcal{T}$. This benefit satisfies the condition $G_{it} \leq G_{i,t-1}, \forall i \in \mathcal{I}, t \in \mathcal{T} \setminus \{1\}$, ensuring that earlier service is always preferred within the planning horizon. Table 1 presents the nomenclature used in our mathematical optimization model, defining the sets, parameters, and decision variables to formulate the problem.

Definition 1 (TRSP-STSU)**.** *The Technician Routing and Scheduling Problem with Skills and Time-Sensitive Returns under Uncertainty (TRSP-STSU) is defined as the problem of scheduling and routing technician teams over a discrete planning horizon $\mathcal{T}$ in order to maximize the total benefit of serviced customers. The problem requires that each prioritized customer $i \in \mathcal{I}^p$ is visited exactly once within the horizon, whereas each optional customer $i \in \mathcal{I}^o$ is visited at most once if time and resources allow. The total working hours for each technician team cannot exceed T^{max} per period. Each team (equivalently vehicle) is staffed with exactly C technicians whose skills must satisfy the following feasibility conditions: for every serviced customer $i \in \mathcal{I}$, and for additive skills $m \in \mathcal{M}^{add}$, the combined team level must meet the requirement R'_{im}, and for qualification skills $m \in \mathcal{M}^{cov}$, at least one technician on the team must individually satisfy the requirement (e.g., $R_{lm} \geq R'_{im}$). Every vehicle departs from and returns to the depot once per period, boards exactly C technicians, and must serve at least one customer during that period. Furthermore, travel and service times $\tilde{\tau}_{ij}$ and $\tilde{S}_i$ are subject to uncertainty. Finally, the benefit G_{it} is non-increasing in time, i.e., $G_{it} \leq G_{i,t-1}$.*

The use of a team-based skill requirement reflects the operational practice of Hydro-Québec, where service tasks are typically performed by multi-technician crews rather than by a single individual. In this setting, task feasibility depends on the collective competencies of the crew. Some requirements are naturally additive (e.g., pooled proficiency or manpower-based capability), while others correspond to certifications or safety qualifications that must be held by at least one crew member. The proposed hybrid structure captures both aspects by combining an additive team-capacity component for $m \in \mathcal{M}^{\text{add}}$ with a coverage-type qualification component for $m \in \mathcal{M}^{\text{cov}}$. While alternative formulations could require that a single designated technician satisfy an entire bundle of require-

Table 1: Summary of Notation

Sets	
$\mathcal{I}^p$	the set of customers that must be visited during the planning horizon $(i, i', j, j' = 1, \dots, I^p)$
$\mathcal{I}^o$	the set of customers that can be visited during the planning horizon $(i, i', j, j' = 1, \dots, I^o)$
$\mathcal{I}$	$\mathcal{I}^o \cup \mathcal{I}^p$
$\mathcal{I}_0$	$\mathcal{I}^o \cup \mathcal{I}^p \cup \{0\}$
$\mathcal{M}$	the set of skill types $(m = 1, \dots, M)$
$\mathcal{M}^{\text{add}}$	the set of additive (team-capacity) skill types
$\mathcal{M}^{\text{cov}}$	the set of qualification (coverage) skill types
$\mathcal{L}$	the set of technicians $(l = 1, \dots, L)$
$\mathcal{T}$	the set of time periods $(t = 1, \dots, T)$
$\mathcal{V}$	the set of vehicles $(v = 1, \dots, V)$
Parameters	
T^{max}	the maximum amount of time each technician team can work during each time period
$\tilde{\tau}_{ij}$	the travel time from node $i \in \mathcal{I}_0$ to node $j \in \mathcal{I}_0$
G_{it}	the benefit of visiting customer $i \in \mathcal{I}$ in period $t \in \mathcal{T}$ $(G_{it} \leq G_{it-1}, \forall i \in \mathcal{I}, t \in \mathcal{T} \backslash \{1\})$
$\tilde{S}_i$	service time of customer $i \in \mathcal{I}$; set $\tilde{S}_0 = 0$ for the depot $(0 \in \mathcal{I}_0)$.
R_{lm}	the skill level of technician $l \in \mathcal{L}$ for skill type $m \in \mathcal{M}$
R'_{im}	minimum required level for customer $i \in \mathcal{I}$ and skill type $m \in \mathcal{M}$
Q_{ilm}	iff technician l satisfies the qualification requirement of customer i for coverage skill $m \in \mathcal{M}^{\text{cov}}$
C	number of technicians that must be assigned to a vehicle
Decision Variables	
x_{ijtv}	iff location $j \in \mathcal{I}_0$ is visited after location $i \in \mathcal{I}_0$ in period $t \in \mathcal{T}$ by vehicle $v \in \mathcal{V}$
y_{ltv}	iff technician $l \in \mathcal{L}$ is on vehicle $v \in \mathcal{V}$ in period $t \in \mathcal{T}$
w_{itv}	iff customer $i \in \mathcal{I}$ is visited in period $t \in \mathcal{T}$ by vehicle $v \in \mathcal{V}$

ments, we do not consider such full "single-technician dominance" constraints in this study, as they are not representative of the operational protocols observed in the collaborating utility. Extending the model to accommodate lead-technician dominance is conceptually straightforward (e.g., via additional assignment variables) and is left for future research.

Definition 2 (Time-Feasible Routing Set)**.** *Let $\bar{\tau}_{ij}$ denote the estimated travel time from i to j and $\bar{S}_i$ the estimated service time at location i (with $\bar{S}_0 = 0$ for the depot). The set $\mathcal{X}_{Dur}$ contains all routing and visit plans (x, w) that respect the maximum allowable working duration T^{max} for every vehicle in every period:*

$$\mathcal{X}_{Dur} = \left\{ (x, w) \;\middle|\; \sum_{i \in \mathcal{I}_0} \sum_{j \in \mathcal{I}_0, j \neq i} \bar{\tau}_{ij} \, x_{ijtv} \;+\; \sum_{i \in \mathcal{I}} \bar{S}_i \, w_{itv} \leq T^{max}, \; \forall t \in \mathcal{T}, \; v \in \mathcal{V} \right\}.$$

3.1 Deterministic Formulation of the TRSP-STSU

This first model assumes that all parameters are known with certainty and it is formulated as follows:

$$\text{Maximize} \quad \sum_{i \in \mathcal{I}_0} \sum_{j \in \mathcal{I}, i \neq j} \sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}} G_{jt} \, x_{ijtv} \tag{1}$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{L}} y_{ltv} = C \qquad \forall v \in \mathcal{V}, \; t \in \mathcal{T} \tag{2}$$

$$\sum_{v \in \mathcal{V}} y_{ltv} = 1 \qquad \forall l \in \mathcal{L},\ t \in \mathcal{T} \tag{3}$$

$$\sum_{i \in \mathcal{I}_0, i \neq j} \sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}} x_{ijtv} = 1 \qquad \forall j \in \mathcal{I}^p \tag{4}$$

$$\sum_{i \in \mathcal{I}_0, i \neq j} \sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}} x_{ijtv} \leq 1 \qquad \forall j \in \mathcal{I}^o \tag{5}$$

$$\sum_{i \in \mathcal{I}_0, i \neq j} x_{ijtv} = \sum_{i \in \mathcal{I}_0, i \neq j} x_{jitv} \qquad \forall j \in \mathcal{I}_0,\ t \in \mathcal{T},\ v \in \mathcal{V} \tag{6}$$

$$\sum_{j \in \mathcal{I}} x_{0jtv} = 1 \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V} \tag{7}$$

$$\sum_{i \in \mathcal{I}} x_{i0tv} = 1 \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V} \tag{8}$$

$$w_{itv} = \sum_{j \in \mathcal{I}_0, j \neq i} x_{jitv} \qquad \forall i \in \mathcal{I},\ t \in \mathcal{T},\ v \in \mathcal{V} \tag{9}$$

$$\sum_{i \in \mathcal{N}} \sum_{\substack{j \in \mathcal{N} \\ j \neq i}} x_{ijtv} \leq |\mathcal{N}| - 1 \qquad \forall \mathcal{N} \in \mathcal{S}_{tv}, |\mathcal{N}| \geq 2,\ t \in \mathcal{T},\ v \in \mathcal{V} \tag{10}$$

$$\sum_{l \in \mathcal{L}} R_{lm}\, y_{ltv} \geq R'_{im}\, w_{itv} \qquad \forall i \in \mathcal{I},\ m \in \mathcal{M}^{\text{add}},\ t \in \mathcal{T},\ v \in \mathcal{V} \tag{11}$$

$$\sum_{l \in \mathcal{L}} Q_{ilm}\, y_{ltv} \geq w_{itv} \qquad \forall i \in \mathcal{I},\ m \in \mathcal{M}^{\text{cov}},\ t \in \mathcal{T},\ v \in \mathcal{V} \tag{12}$$

$$(x, w) \in \mathcal{X}_{Dur} \tag{13}$$

$$x_{ijtv},\ y_{ltv},\ w_{itv} \in \{0, 1\} \qquad \forall i, j \in \mathcal{I}_0,\ l \in \mathcal{L},\ t \in \mathcal{T},\ v \in \mathcal{V}. \tag{14}$$

The objective function (1) maximizes the total benefit obtained from servicing customers over the planning horizon. Constraints (2) ensure that exactly C technicians are assigned to each vehicle in every time period, while constraints (3) guarantee that each technician is assigned to exactly one vehicle per period. Customer visitation requirements are governed by constraints (4) and (5): the former enforces that each prioritized customer is visited exactly once over the entire planning horizon, whereas the latter allows optional customers to be visited at most once. The flow conservation constraints (6), together with the depot departure and return constraints (7) and (8), ensure route continuity and feasibility for each vehicle in every period. The visit-indicator linking constraints (9) define the binary variable w_{itv} as an exact indicator of whether customer i is visited by vehicle v in period t, by equating it to the corresponding incoming routing flow.

Using this definition, the Subtour Elimination Constraints (SEC) (10) prevent the formation of disconnected cycles. SECs are exponentially many. We therefore add them as lazy constraints and separate only violated subtours found in incumbent integer solutions during the BC search. They are initially relaxed and dynamically separated within a BC framework. Specifically, the set $\mathcal{S}_{tv}$ represents the family of violated subtours identified by the solver during the optimization process. For any identified subset $\mathcal{N}$ forming a disconnected cycle, the corresponding constraint is generated and added to the model, ensuring that the inequality holds only for subsets of visited nodes.

Constraints (11) and (12) enforce hybrid team-skill feasibility. For additive skill types $m \in \mathcal{M}^{\text{add}}$, (11) requires that the aggregate skill levels of technicians assigned to a vehicle meet the minimum requirements of any customer served. For qualification skill types $m \in \mathcal{M}^{\text{cov}}$, (12) requires that at least one technician on the team satisfies the customer-specific threshold, encoded by Q_{ilm}. The time-feasibility set (13) enforces an upper bound on the total travel and service time incurred by each vehicle in every period, ensuring that any constructed route remains feasible within the maximum allowable working duration. Finally, the binary nature of the routing, assignment, and visit-indicator variables is enforced by constraints (14).

Remark 1. *The model assumes that all technicians are dispatched in every time period, as enforced by constraints* (2)–(8)*, which require each vehicle to depart from and return to the depot in each period with a full team. This assumption is well aligned with the operational context of Hydro-Québec, where technician crews are scheduled on fixed shifts and vehicles are typically rolled out once a crew is assigned, even if the workload in a given period is limited. From an operational standpoint, this reflects labor contracts, safety regulations, and readiness requirements. Nevertheless, this model can present a more general case issue without altering the structure of the model, as the formulation can be extended by introducing dummy customer nodes with zero benefit, i.e., $G_{it} = 0$, and negligible service and travel times. These dummy nodes allow vehicles to remain idle while preserving route feasibility and crew assignment constraints. As a result, the model retains full flexibility to decide whether a vehicle performs productive service or effectively remains inactive in a given period, without relaxing the dispatch or assignment constraints.*

3.2 Chance-Constrained Formulation

We now incorporate uncertainty in travel times and service durations. We assume that the arc travel times $\tilde{\tau}_{ij}$ and node service times $\tilde{S}_i$ are independent and normally distributed with known means $\bar{\tau}_{ij}$, $\bar{S}_i$ and variances $\sigma^2_{\tau_{ij}}$, $\sigma^2_{S_i}$, respectively. Independence between travel and service components is motivated by their different driving factors: travel times are largely shaped by external conditions (e.g., traffic and congestion), whereas service durations reflect job- and customer-specific characteristics (e.g., task complexity and equipment condition). In addition, in many field-service settings, service durations are sufficiently long that dependencies across successive travel arcs are reduced relative to short-cycle distribution contexts.

For each period $t \in \mathcal{T}$ and vehicle $v \in \mathcal{V}$, we require that the total on-duty time does not exceed the per-period limit T^{max} with probability at least α:

$$\Pr\left(\sum_{i \in \mathcal{I}_0} \sum_{\substack{j \in \mathcal{I}_0 \\ j \neq i}} \tilde{\tau}_{ij}\, x_{ijtv} \;+\; \sum_{i \in \mathcal{I}} \tilde{S}_i\, w_{itv} \;\leq\; T^{max} \right) \geq \alpha, \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V}. \tag{15}$$

Here, the travel time accumulates over selected arcs via x_{ijtv}, and the service time accumulates once per visited customer via the visit indicator w_{itv}.

Under the normality and independence assumptions, the left-hand side of (15) is normally distributed. Let $z_\alpha := \Phi^{-1}(\alpha)$ denote the standard normal quantile. The chance constraint admits the deterministic (conic) equivalent

$$\sum_{i\in\mathcal{I}_0}\sum_{\substack{j\in\mathcal{I}_0\\ j\neq i}} \bar{\tau}_{ij}\, x_{ijtv} \;+\; \sum_{i\in\mathcal{I}} \bar{S}_i\, w_{itv} \;+\; z_\alpha \sqrt{\sum_{i\in\mathcal{I}_0}\sum_{\substack{j\in\mathcal{I}_0\\ j\neq i}} \sigma^2_{\tau_{ij}}\, x_{ijtv} \;+\; \sum_{i\in\mathcal{I}} \sigma^2_{S_i}\, w_{itv}} \;\le\; T^{max}, \quad \forall t\in\mathcal{T},\ v\in\mathcal{V}. \tag{16}$$

To express (16) in solver-ready second-order cone (SOC) form, we introduce an auxiliary variable $z_{tv} \geq 0$ and impose

$$\sum_{i\in\mathcal{I}_0}\sum_{\substack{j\in\mathcal{I}_0\\ j\neq i}} \bar{\tau}_{ij}\, x_{ijtv} \;+\; \sum_{i\in\mathcal{I}} \bar{S}_i\, w_{itv} \;+\; z_\alpha\, z_{tv} \;\le\; T^{max}, \qquad \forall t\in\mathcal{T},\ v\in\mathcal{V}, \tag{17}$$

$$\left\| \left(\sqrt{\sigma^2_{\tau_{ij}}}\, x_{ijtv}\right)_{(i,j):\, i\neq j} \;\oplus\; \left(\sqrt{\sigma^2_{S_i}}\, w_{itv}\right)_{i\in\mathcal{I}} \right\|_2 \;\le\; z_{tv}, \qquad \forall t\in\mathcal{T},\ v\in\mathcal{V}, \tag{18}$$

$$z_{tv} \geq 0, \qquad \forall t\in\mathcal{T},\ v\in\mathcal{V}. \tag{19}$$

Constraint (18) is a standard SOC constraint and is directly supported by commercial Mixed-Integer Second-Order Cone Program (MISOCP) solvers. Together, (17)–(19) enforce that the expected route duration plus a safety margin proportional to the standard deviation of total time (scaled by z_α) does not exceed T^{max}. Hence, replacing the deterministic duration set $\mathcal{X}_{\text{Dur}}$ with (17)–(19) yields a MISOCP.

Definition 3 (Stochastic Time-Feasible Routing Set)**.** *The set $\mathcal{X}_{Dur}^{stoch}$ contains all routing and visit plans (x, w) that satisfy the chance-constrained time-feasibility condition under normally distributed travel and service times.*

$$\mathcal{X}_{Dur}^{stoch} = \left\{ (x,w) \;\middle|\; \Pr\!\left(\sum_{i\in\mathcal{I}_0}\sum_{\substack{j\in\mathcal{I}_0\\ j\neq i}} \tilde{\tau}_{ij}\, x_{ijtv} + \sum_{i\in\mathcal{I}} \tilde{S}_i\, w_{itv} \le T^{max} \right) \ge \alpha,\ \forall t\in\mathcal{T},\ v\in\mathcal{V} \right\}.$$

Constraints (17) enforce that the expected total time plus a safety margin $z_\alpha z_{tv}$ does not exceed T^{max}, while (18) defines z_{tv} as an upper bound on the standard deviation of the total time. Together, these constraints provide a solver-ready conic reformulation, so that the resulting stochastic model is a MISOCP.

$$\begin{aligned} \text{Maximize} \quad & \sum_{i\in\mathcal{I}_0}\sum_{j\in\mathcal{I}, i\neq j}\sum_{t\in\mathcal{T}}\sum_{v\in\mathcal{V}} G_{jt} x_{ijtv} \\ \text{s.t.} \quad & (2) - (12), (14) \\ & (x,w) \in \mathcal{X}_{\text{Dur}}^{\text{stoch}}. \end{aligned} \tag{20}$$

Remark 2. *Our baseline model assumes independence across all travel-time variables and across all service-time variables, as well as independence between the two categories. However, the chance-constrained formulation is readily extended to accommodate within-category correlations. In this setting, travel times across arcs may share common stochastic factors (e.g., traffic conditions), and service durations across customers may be correlated (e.g., technician effects or customer complexity). A common approach is to represent each* *uncertain parameter as a linear combination of independent factors, following Jaillet et al. (2016). For example,*

$$\tilde{\tau}_{ij} \;=\; \bar{\tau}_{ij} \;+\; \sum_{b=1}^{B} \tau_{ij}^{b} c_b, \qquad \tilde{S}_i \;=\; \bar{S}_i \;+\; \sum_{b=1}^{B'} S_i^{b} c'_b,$$

where (c_b) *and* (c'_b) *are independent zero-mean factors capturing shared traffic or shared service-complexity effects.*

If such within-category correlations are incorporated, the chance constraint remains second-order cone representable. Let x denote the vector of active arcs and w the vector of customer visits, and let Σ *be the block-diagonal covariance matrix capturing correlations among* $\tilde{\tau}_{ij}$ *and among* $\tilde{S}_i$ *(but not across the two categories). The deterministic equivalent simply becomes*

$$\bar{t}^{\top} x \;+\; \bar{s}^{\top} w \;+\; \Phi^{-1}(\alpha)\,\left\|\Sigma^{1/2}[x; w]\right\|_2 \;\le\; T^{max},$$

which is the standard SOC form for a normally distributed linear combination with covariance Σ.

We do not use this correlated version in our computational study, but the extension is fully compatible with our overall MISOCP framework.

4 Solution Approaches

In this section, we present the methodology employed to solve the problem defined in Section 3. We propose a tailored LBBD algorithm as our primary solution approach. The LBBD is particularly well-suited for problems with intricate combinatorial structures, enabling a natural decomposition into assignment and routing sub-problems. For completeness, an alternative method (BC) is also utilized. This algorithm serves as a benchmark for evaluating the LBBD's performance in terms of scalability, solution quality, and robustness.

4.1 Logic-Based Benders Decomposition (LBBD)

In LBBD, we decompose the problem into a master problem to determine the primary assignment of technicians and customers to vehicles and a set of sub-problems to evaluate the feasibility of the master problem's solutions under the specified constraints. This iterative process involves generating Benders cuts, including no-good cuts, which are added to the master problem to eliminate infeasible solutions and guide the search toward optimality.

4.1.1 Master Problem

The master problem assigns customers and technicians to vehicles while ensuring feasibility concerning coverage, availability, and hybrid skill requirements:

$$\text{Maximize} \quad \sum_{i\in\mathcal{I}} \sum_{t\in\mathcal{T}} \sum_{v\in\mathcal{V}} G_{it} w_{itv}$$

$$\text{s.t.} \quad \sum_{l\in\mathcal{L}} y_{ltv} = C \qquad \forall v \in \mathcal{V}, t \in \mathcal{T} \tag{21}$$

$$\sum_{v \in \mathcal{V}} y_{ltv} = 1 \qquad \forall l \in \mathcal{L}, t \in \mathcal{T} \tag{22}$$

$$\sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}} w_{itv} = 1 \qquad \forall i \in \mathcal{I}^p \tag{23}$$

$$\sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}} w_{itv} \leq 1 \qquad \forall i \in \mathcal{I}^o \tag{24}$$

$$\sum_{i \in \mathcal{I}} w_{itv} \geq 1 \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V} \tag{25}$$

$$\sum_{l \in \mathcal{L}} R_{lm} y_{ltv} \geq R'_{im} w_{itv} \qquad \forall i \in \mathcal{I},\ m \in \mathcal{M}^{\text{add}},\ t \in \mathcal{T},\ v \in \mathcal{V} \tag{26}$$

$$\sum_{l \in \mathcal{L}} Q_{ilm}\, y_{ltv} \geq w_{itv} \qquad \forall i \in \mathcal{I},\ m \in \mathcal{M}^{\text{cov}},\ t \in \mathcal{T},\ v \in \mathcal{V} \tag{27}$$

$$(w, y) \in \mathcal{F}_{\text{LBBD}}, \tag{28}$$

$$y_{ltv}, w_{itv} \in \{0, 1\} \qquad \forall i \in \mathcal{I}, l \in \mathcal{L}, t \in \mathcal{T}, v \in \mathcal{V}. \tag{29}$$

Constraints (21)–(24) enforce vehicle capacity, technician assignment, and customer coverage rules. Constraint (25) ensures that each vehicle serves at least one customer in every period, maintaining consistency with the routing structure of the overall TRSP-STSU formulation. Constraints (26)–(27) enforce hybrid skill feasibility: for additive skill types $m \in \mathcal{M}^{\text{add}}$, (26) requires that the aggregate skill level of the team assigned to vehicle v meets the customer requirement, while for qualification skill types $m \in \mathcal{M}^{\text{cov}}$, (27) requires that at least one assigned technician is individually qualified for customer i (captured by Q_{ilm}).

Note that the master problem does not include explicit routing variables. Instead, routing decisions are deferred to the sub-problems: for each vehicle v and time period t, the set $\mathcal{A}_{tv} = \{i \in \mathcal{I} : w_{itv} = 1\}$ is passed to the corresponding sub-problem. Constraint (28) represents the collection of LBBD feasibility cuts that are dynamically added during the solution process.

4.1.2 Benders Cuts and No-Good Cuts

Building upon the route-based formulation, our LBBD decomposition introduces feasibility cuts to iteratively refine the master solution. These cuts fall into two categories: (i) *mandatory feasibility cuts*, which ensure correctness by eliminating master assignments that cannot be extended to a feasible route, and (ii) *optional enhancement cuts*, which are not required for correctness but substantially improve computational performance by strengthening the master and accelerating convergence.

In our implementation, the master problem is solved as an MIP and feasibility cuts are added dynamically as lazy constraints. For each vehicle $v \in \mathcal{V}$ and period $t \in \mathcal{T}$, the master induces the customer set

$$\mathcal{A}_{tv} \ := \ \{\, i \in \mathcal{I} : w_{itv} = 1 \,\},$$

which is passed to the routing subproblem. The routing subproblem checks whether $\mathcal{A}_{tv} \cup \{0\}$ admits a depot-starting tour satisfying the per-period duration budget T^{max} in the deterministic case, or the chance-constrained time budget in the stochastic case. Whenever the subproblem is infeasible, we generate cuts that prevent the master from repeating the same (or any provably infeasible) customer grouping.

Global customer-set no-good cuts If the routing subproblem is infeasible for a customer set $\mathcal{A}_{tv}$, then this set cannot be assigned to *any* single route because the duration parameters and the budget T^{max} are identical across periods and vehicles in our setting (and, in the chance-constrained model, the distributional parameters and confidence level α are also stationary). We therefore add the following *global* no-good cut:

$$\sum_{i \in \mathcal{A}_{tv}} w_{it'v'} \;\leq\; |\mathcal{A}_{tv}| - 1 \qquad \forall\, t' \in \mathcal{T},\ \forall\, v' \in \mathcal{V}. \tag{30}$$

Cut (30) eliminates the simultaneous assignment of all customers in $\mathcal{A}_{tv}$ to the same vehicle in any period, while preserving flexibility for alternative allocations (e.g., splitting $\mathcal{A}_{tv}$ across different routes or periods). These cuts are mandatory for correctness in their standard (local) form for the incumbent pair (t, v). Under our stationarity assumption (time parameters and T^{max} identical across all periods and vehicles), we further generalize them over all (t', v') as in (30), which is a valid strengthening.

Minimal conflict-set (cover) cuts The global no-good cut (30) excludes only the *entire* infeasible set $\mathcal{A}_{tv}$. In practice, infeasibility is often driven by a smaller subset of customers that already violates the time budget irrespective of sequencing. To strengthen the master earlier and improve reuse of cuts, we derive a *conflict set* $\mathcal{K} \subseteq \mathcal{A}_{tv}$ such that any route covering all customers in $\mathcal{K}$ is time-infeasible, independent of visit order.

For the deterministic model, define the sequence-independent lower bound

$$\delta_0^{\mu} \;:=\; \min_{i \in \mathcal{I}} \bar{\tau}_{0i}, \qquad \underline{c}_i^{\mu} \;:=\; \bar{S}_i \;+\; \min_{j \in \mathcal{I}_0 \setminus \{i\}} \bar{\tau}_{ij}, \qquad \mathcal{I}_0 := \mathcal{I} \cup \{0\}.$$

Then, for any nonempty $\mathcal{S} \subseteq \mathcal{I}$, $LB_{\mu}(\mathcal{S}) \;:=\; \delta_0^{\mu} \;+\; \sum_{i \in \mathcal{S}} \underline{c}_i^{\mu}$ is a valid lower bound on the duration of any depot-starting tour visiting all customers in $\mathcal{S}$. Whenever a subset $\mathcal{K} \subseteq \mathcal{A}_{tv}$ satisfies $LB_{\mu}(\mathcal{K}) > T^{max}$, no feasible route can cover all customers in $\mathcal{K}$, regardless of sequencing, and we add the global cover cut

$$\sum_{i \in \mathcal{K}} w_{it'v'} \;\leq\; |\mathcal{K}| - 1 \qquad \forall\, t' \in \mathcal{T},\ \forall\, v' \in \mathcal{V}. \tag{31}$$

In our implementation, $\mathcal{K}$ is obtained by shrinking $\mathcal{A}_{tv}$ via a greedy deletion procedure until the bound becomes tight (i.e., no further customer can be removed while keeping $LB_{\mu}(\mathcal{K}) > T^{max}$). This typically yields small, reusable conflict sets and substantially reduces repeated infeasibility checks.

Chance-constrained conflict sets For the chance-constrained model, the routing feasibility uses a mean–safety-margin form. We therefore strengthen the conflict detection with a bound that incorporates both mean and variance. Define the variance proxies

$$\delta_0^{\sigma^2} := \min_{i \in \mathcal{I}} \sigma^2_{\tau_{0i}}, \qquad \underline{v}_i := \sigma^2_{S_i} + \min_{j \in \mathcal{I}_0 \setminus \{i\}} \sigma^2_{\tau_{ij}}.$$

Then, for any nonempty $\mathcal{S} \subseteq \mathcal{I}$, the bound

$$LB_{\mathrm{cc}}(\mathcal{S}) \ := \ \delta_0^{\mu} + \sum_{i \in \mathcal{S}} \underline{c}_i^{\mu} \ + \ z_\alpha \sqrt{\delta_0^{\sigma^2} + \sum_{i \in \mathcal{S}} \underline{v}_i}$$

is a valid sequence-independent lower bound on the chance-constrained tour duration. Whenever a subset $\mathcal{K} \subseteq \mathcal{A}_{tv}$ satisfies $LB_{\mathrm{cc}}(\mathcal{K}) > T^{max}$, we add the same global cover cut (31). This provides strong, model-consistent infeasibility certificates and improves the effectiveness of early pruning under uncertainty.

Fractional cover user cuts (LP-relaxation strengthening) In addition to lazy feasibility cuts at integer incumbents, we also strengthen the master relaxation by separating violated cover inequalities at fractional solutions. Concretely, at selected nodes of the branch-and-bound tree, we inspect the current LP relaxation values of the assignment variables and greedily construct a candidate customer subset C^{cov} with large fractional mass. If the sequence-independent bound certifies that C^{cov} is time-infeasible—i.e., if $LB_{\mu}(C^{\mathrm{cov}}) > T^{max}$ in the deterministic case or $LB_{\mathrm{cc}}(C^{\mathrm{cov}}) > T^{max}$ in the chance-constrained case—we add the valid cover inequality

$$\sum_{i \in C^{\mathrm{cov}}} w_{itv} \leq |C^{\mathrm{cov}}| - 1 \qquad \forall\, t \in \mathcal{T},\ v \in \mathcal{V},$$

as a *user cut*. Since these cuts are derived from the same sequence-independent infeasibility certificates as in the integer-feasibility separation, they do not exclude any integer-feasible solution and primarily serve to tighten the LP bounds and accelerate convergence.

Remark 3. *In the present study, routing feasibility depends only on the assigned customer set $\mathcal{A}_{tv}$ and the duration budget, and is independent of the specific technician team assigned to the vehicle. Consequently, feasibility cuts are generated in the w-space as above. If future extensions introduce technician-dependent service times or other crew-dependent routing features, analogous configuration cuts can be added over (w, y) to exclude infeasible customer–crew patterns.*

4.1.3 Valid Inequalities

We strengthen both the mathematical formulations (MIP/MISOCP) and the LBBD master with a family of *valid inequalities* that preserve feasibility and optimality while substantially reducing the search space. These inequalities (i) preclude provably time-infeasible visit combinations, (ii) cap overpacked daily routes early, and (iii) mitigate vehicle-label symmetry. They are enforced for every period $t \in \mathcal{T}$ and vehicle $v \in \mathcal{V}$. Throughout the paper, we assume that travel times satisfy the triangle inequality, i.e., $\bar{\tau}_{ik} \leq \bar{\tau}_{ij} + \bar{\tau}_{jk}$ for all $i, j, k \in \mathcal{I}_0$ (and analogously for any deterministic travel-time surrogate used in cut generation), and that service times are nonnegative. This standard metric assumption ensures that any tour visiting a given customer subset cannot be shorter than the best depot tour over any of its sub-subsets, which justifies the pairwise incompatibility construction and related clique inequalities, as well as arc-elimination preprocessing rules used in the monolithic BC/MIP.

Deterministic vs. stochastic bounding convention. All inequalities below are generated using *deterministic surrogates* of travel and service times. In the deterministic model, we use mean quantities $\bar{\tau}_{ij}$ and $\bar{S}_i$. In the stochastic (chance-constrained) model with confidence level α, we optionally strengthen these bounds by incorporating a conservative risk margin based on the variances $\sigma^2_{\tau_{ij}}$ and $\sigma^2_{S_i}$. Importantly, throughout the stochastic model we assume $\alpha \in [0.5, 1)$, hence $z_\alpha \geq 0$. Therefore, any inequality derived from mean-only lower bounds remains valid under the chance constraint, since feasibility requires a mean term plus a nonnegative safety margin.

Unified per-customer lower bounds. Let $\mathcal{I}_0 := \mathcal{I} \cup \{0\}$ denote the customer set augmented with the depot 0, the quantity δ_0^μ lower-bounds the first depot-to-customer leg in any nonempty route, and $\underline{c}_i^\mu$ lower-bounds service at i plus the outgoing travel arc from i (possibly the return to the depot for the last visited customer). This construction avoids double-counting depot legs when the minimizer in $\min_{j \in \mathcal{I}_0 \setminus \{i\}} \bar{\tau}_{ij}$ happens to be $j = 0$. A *chance-consistent* lower bound for any nonempty subset $C^{\text{cov}} \subseteq \mathcal{I}$ is:

$$LB_{\text{cc}}(C^{\text{cov}}) := \delta_0^\mu + \sum_{i \in C^{\text{cov}}} \underline{c}_i^\mu + z_\alpha \sqrt{\delta_0^{\sigma^2} + \sum_{i \in C^{\text{cov}}} \underline{v}_i}. \tag{32}$$

Using mean-only bounds corresponds to setting the last term to zero in (32).

Duration cover inequalities (generalized packing). For any nonempty subset $C^{\text{cov}} \subseteq \mathcal{I}$, the expression $\delta_0^\mu + \sum_{i \in C^{\text{cov}}} \underline{c}_i^\mu$ is a valid sequence-independent lower bound on the duration of any depot-starting tour that visits all customers in C^{cov}. In the chance-constrained model, the stronger bound $LB_{\text{cc}}(C^{\text{cov}})$ in (32) is also sequence-independent and conservative.

Whenever a subset C^{cov} violates the (deterministic or chance-consistent) duration bound, i.e.,

$$\delta_0^\mu + \sum_{i \in C^{\text{cov}}} \underline{c}_i^\mu > T^{max} \quad \text{(deterministic)} \qquad \text{or} \qquad LB_{\text{cc}}(C^{\text{cov}}) > T^{max} \quad \text{(chance-consistent)},$$

we add the cover cut

$$\sum_{i \in C^{\text{cov}}} w_{itv} \;\leq\; |C^{\text{cov}}| - 1 \qquad \forall\, t \in \mathcal{T},\; v \in \mathcal{V}. \tag{33}$$

Inequality (33) is a knapsack cover inequality; pairwise and triple infeasibilities are special cases (Balas, 1975; Nemhauser and Wolsey, 1988).

Pairwise time–incompatibility (size-2 specialization). For any distinct $i, j \in \mathcal{I}$, consider the two possible depot tours $(0, i, j, 0)$ and $(0, j, i, 0)$. If both tours violate the time budget under the chosen bound (mean-only or chance-consistent), then i and j cannot be assigned to the same (t, v):

$$w_{itv} + w_{jtv} \leq 1 \qquad \forall\, t \in \mathcal{T},\; v \in \mathcal{V}. \tag{34}$$

This is exactly the $|C^{\text{cov}}| = 2$ case of (33) and is commonly used to prune impossible pairs in routing problems with duration limits (Toth and Vigo, 2014; Savelsbergh, 1992).

Triple infeasibility (size-3 specialization). As a small-cardinality instance of (33), if every depot-start/finish permutation over $\{a, b, c\}$ violates the time budget (under the chosen bound), then

$$w_{atv} + w_{btv} + w_{ctv} \leq 2 \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V}. \tag{35}$$

Such low-cardinality cover inequalities are inexpensive to detect and effective early in the search (Balas, 1975).

Clique inequalities on the incompatibility graph. Build an incompatibility graph on $\mathcal{I}$ with an edge (i, j) whenever the pair (i, j) is time-infeasible according to the test used to generate (34) (mean-only or chance-consistent). For any clique C^{cov} in this graph, we add

$$\sum_{i \in C^{\text{cov}}} w_{itv} \leq 1 \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V}. \tag{36}$$

Clique inequalities dominate the corresponding set of pairwise constraints and tighten the relaxation without excluding feasible integer solutions (Nemhauser and Wolsey, 1988; Padberg and Rinaldi, 1991). In practice, we extract maximal cliques via a deterministic heuristic and add the resulting inequalities once before optimization.

Route cardinality upper bound. Using δ_0^{μ} and $\underline{c}_i^{\mu}$, we derive a conservative upper bound on the number of customers that can be visited in any (t, v):

$$U := \left\lfloor \frac{T^{max} - \delta_0^{\mu}}{\min_{i \in \mathcal{I}} \underline{c}_i^{\mu}} \right\rfloor, \qquad U \geq 1.$$

We then enforce

$$\sum_{i \in \mathcal{I}} w_{itv} \leq U \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V}. \tag{37}$$

If desired in the stochastic model, U can be tightened by replacing $\underline{c}_i^{\mu}$ with a chance-consistent per-visit surrogate; we keep (37) mean-based to avoid excessive conservatism.

Strengthened duration lower bound (assignment-based). Independently of visit sequencing, any feasible tour covering $\{i : w_{itv} = 1\}$ must spend at least one depot-to-customer leg plus one unavoidable chunk per selected customer. We therefore add

$$\delta_0^{\mu} + \sum_{i \in \mathcal{I}} \underline{c}_i^{\mu} w_{itv} \leq T^{max} \qquad \forall t \in \mathcal{T},\ v \in \mathcal{V}, \tag{38}$$

which removes assignment patterns that are sequence-independently infeasible. In the chance-constrained model, (38) remains valid because it is implied by the corresponding SOC time budget (mean term plus nonnegative safety margin). If additional strength is desired, one may replace the left-hand side by the chance-consistent counterpart obtained from (32) with $C^{\text{cov}} = \{i : w_{itv} = 1\}$, at the expense of introducing a conic term in the master.

Linear chance-consistent proxy (exact master strengthening under uncertainty). In the chance-constrained model, feasibility requires controlling both the expected route duration and a nonnegative safety margin. The exact safety margin in (16) depends on routing variables x through the term $\sqrt{\sum_{(i,j):\,i\neq j} \sigma^2_{\tau_{ij}} x_{ijtv} \; + \; \sum_{i\in\mathcal{I}} \sigma^2_{S_i} w_{itv}}$, which is not available in the LBBD master. To strengthen the master without introducing routing variables or conic constraints, we impose an *exactness-preserving* linear screening inequality expressed only in the assignment variables w. The key requirement is that any additional master inequality must be a *necessary* condition for chance-feasible routing (i.e., it cannot exclude any chance-feasible assignment).

Let U be the route cardinality bound enforced by (37), and define $k := \sum_{i\in\mathcal{I}} w_{itv}$ so that $k \le U$. Any feasible route visiting the selected customers therefore satisfies the variance lower bound

$$\sum_{i\in\mathcal{I}_0} \sum_{\substack{j\in\mathcal{I}_0\\ j\neq i}} \sigma^2_{\tau_{ij}} x_{ijtv} \; + \; \sum_{i\in\mathcal{I}} \sigma^2_{S_i} w_{itv} \;\ge\; \delta_0^{\sigma^2} \; + \; \sum_{i\in\mathcal{I}} \underline{v}_i\, w_{itv}. \tag{39}$$

Taking square roots preserves the inequality (both sides are nonnegative), hence the true standard deviation term is at least $\sqrt{\delta_0^{\sigma^2} + \sum_{i\in\mathcal{I}} \underline{v}_i\, w_{itv}}$. However, this expression is not linear in w and cannot be enforced directly in the master.

To retain a purely linear master *without sacrificing exactness*, we replace the nonlinear term by a linear *lower* bound (underestimator) that is valid for all assignments satisfying $k \le U$. Specifically, for any nonnegative numbers $\{a_i\}$ and any integer $k \ge 1$, the Cauchy–Schwarz inequality implies

$$\sum_{i=1}^{k} \sqrt{a_i} \;\le\; \sqrt{k}\sqrt{\sum_{i=1}^{k} a_i} \qquad \Longrightarrow \qquad \sqrt{\sum_{i=1}^{k} a_i} \;\ge\; \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \sqrt{a_i}.$$

Applying this inequality to the selected customers (with $a_i = \underline{v}_i$ for those with $w_{itv} = 1$) and using $k \le U$ yields the linear, assignment-only bound

$$\sqrt{\delta_0^{\sigma^2} + \sum_{i\in\mathcal{I}} \underline{v}_i\, w_{itv}} \;\ge\; \frac{\lambda}{\sqrt{U}} \sum_{i\in\mathcal{I}} \sqrt{\underline{v}_i}\, w_{itv}, \qquad \forall t\in\mathcal{T},\; v\in\mathcal{V}, \tag{40}$$

where $\lambda \in [0,1]$ is a tunable screening parameter. Since the right-hand side of (40) is a *lower bound* on the true safety margin for any $k \le U$, substituting it into the chance constraint produces a *necessary* master-level screening condition and therefore preserves LBBD exactness.

Combining (40) with the sequence-independent mean lower bound $\delta_0^{\mu} + \sum_{i\in\mathcal{I}} \underline{c}_i^{\mu} w_{itv}$ yields the following purely linear master strengthening inequality:

$$\delta_0^{\mu} \; + \; \sum_{i\in\mathcal{I}} \underline{c}_i^{\mu}\, w_{itv} \; + \; z_\alpha \frac{\lambda}{\sqrt{U}} \sum_{i\in\mathcal{I}} \sqrt{\underline{v}_i}\, w_{itv} \;\le\; T^{max} \qquad \forall t\in\mathcal{T},\; v\in\mathcal{V}. \tag{41}$$

Inequality (41) is a lightweight *necessary* condition for chance-feasible routing: it uses a sequence-independent lower bound on the expected duration and an exactness-preserving linear lower bound on the safety margin. Because (41) is linear in w, it can be imposed directly in the LBBD master to strengthen the relaxation while preserving a purely linear master formulation. The parameter λ provides a practical calibration knob: $\lambda = 1$ yields the strongest screening that remains exactness-preserving under the cardinality bound $k \le U$, while smaller values weaken the screen and may improve numerical robustness.

Vehicle-load ordering (symmetry breaking). For identical vehicles, we impose a non-increasing load profile across labels:

$$\sum_{i\in\mathcal{I}} w_{itv} \geq \sum_{i\in\mathcal{I}} w_{i,t,v+1} \qquad \forall\, t\in\mathcal{T},\; v\in\mathcal{V}\setminus\{V\}. \tag{42}$$

This eliminates vehicle-label permutation symmetry and does not remove any feasible solution, provided vehicles are indistinguishable and crews are decision variables (so vehicle labels can be permuted together with the (w, y) assignments) (Sherali and Smith, 2001; Toth and Vigo, 2014).

Implementation note. All inequalities in (33)–(41) are generated once prior to optimization via deterministic, instance-dependent tests. In the stochastic model, we generate incompatibility and cover sets either (i) using mean-only surrogates (always valid, less aggressive) or (ii) using the chance-consistent bound (32) (valid and typically stronger). The linear proxy (41) is enforced only in the chance-constrained runs and relies on the route-cardinality bound (37).

4.1.4 Uniform Preprocessing (model-agnostic)

Before optimization, we apply a light, model-agnostic preprocessing that removes provably infeasible visits and fixes variables without changing the feasible region. All tests use the same mean quantities that define the deterministic duration set $\mathcal{X}_{\text{Dur}}$ (Definition 2): $\bar{\tau}_{ij}$ for travel and $\bar{S}_i$ for service. The resulting reductions are applied identically to every algorithmic variant (BC/MIP, LBBD), ensuring fair comparisons.

(P1) Unserviceable singletons. If even the shortest depot tour that visits customer i alone exceeds the per-period time budget, then i cannot be served by any vehicle in any period. Hence, we fix

$$\bar{\tau}_{0i} + \bar{S}_i + \bar{\tau}_{i0} \;>\; T^{max} \implies w_{itv} = 0 \quad \forall\, t\in\mathcal{T},\, v\in\mathcal{V}.$$

This rule is valid for the deterministic model and remains valid under chance-constrained routing, since any chance constraint enforces a mean term plus a nonnegative safety margin.

(P2) Skill-infeasible customers. Let C denote the team size per vehicle and assume the hybrid skill partition $\mathcal{M} = \mathcal{M}^{\text{add}} \cup \mathcal{M}^{\text{cov}}$, where additive skills are pooled at the team level and qualification skills require at least one individually qualified technician.

(P2a) Additive (team-capacity) skills. For each $m \in \mathcal{M}^{\text{add}}$, define the maximum total skill that any size-C team can provide as

$$\overline{R}_m \;:=\; \max_{\substack{\mathcal{L}'\subseteq\mathcal{L}\\ |\mathcal{L}'|=C}} \sum_{l\in\mathcal{L}'} R_{lm},$$

which is obtained by summing the top-C values of $\{R_{lm} : l \in \mathcal{L}\}$. If customer i requires R'_{im} units of additive skill m and $\overline{R}_m < R'_{im}$, then no feasible team can ever qualify for i, and we fix

$$\exists\, m \in \mathcal{M}^{\text{add}} :\; \overline{R}_m < R'_{im} \implies w_{itv} = 0 \quad \forall\, t\in\mathcal{T},\, v\in\mathcal{V}.$$

(P2b) Qualification (Coverage) Skills.

For each $m \in \mathcal{M}^{\text{cov}}$, let $Q_{ilm} = 1$ if technician l is individually qualified to serve customer i with respect to skill m (i.e., $R_{lm} \geq R'_{im}$), and $Q_{ilm} = 0$ otherwise. If, for at least one required coverage skill, no technician is qualified, then customer i is deemed unserviceable. Formally,

$$\exists\, m \in \mathcal{M}^{\text{cov}} : \sum_{l \in \mathcal{L}} Q_{ilm} = 0 \implies w_{itv} = 0 \quad \forall\, t \in \mathcal{T},\, v \in \mathcal{V}.$$

Rules (P1)–(P2) only eliminate assignments that violate the time-feasibility convention in $\mathcal{X}_{\text{Dur}}$ or the hybrid master skill requirements (Constraints (26)–(27)); therefore, they do not remove any feasible solution of the proposed models. In practice, these filters substantially reduce the search space and uniformly strengthen all methods.

Preprocessing rules (Section 4.1.4) are applied once to fix variables and shrink the instance, whereas valid inequalities (Section 4.1.3) remain in the formulation to tighten the relaxation throughout the branch-and-bound search.

4.1.5 Sub-problem

For each vehicle $v \in \mathcal{V}$ and period $t \in \mathcal{T}$, the routing sub-problem checks whether the customer set induced by the master,

$$\mathcal{A}_{tv} := \{\, i \in \mathcal{I} : w_{itv} = 1 \,\}, \qquad \mathcal{N}_{tv} := \mathcal{A}_{tv} \cup \{0\},$$

can be sequenced into a single depot-starting tour that satisfies the per-period time budget. In the deterministic case, we solve a *minimum-duration TSP* on $\mathcal{N}_{tv}$ and declare the set $\mathcal{A}_{tv}$ *feasible* if and only if the optimal tour duration does not exceed T^{max}.

Deterministic sub-problem (minimum-duration TSP). Let $x_{ij}^{tv} = 1$ if node $j \in \mathcal{N}_{tv}$ is visited immediately after node $i \in \mathcal{N}_{tv}$ in the tour of (t, v), and $x_{ij}^{tv} = 0$ otherwise. The deterministic sub-problem is:

$$(\text{SP}_{tv}^{\text{det}}) \qquad \min \sum_{i \in \mathcal{N}_{tv}} \sum_{\substack{j \in \mathcal{N}_{tv} \\ j \neq i}} \left(\bar{\tau}_{ij} + \bar{S}_i\right) x_{ij}^{tv} \tag{43}$$

$$\text{s.t.} \quad \sum_{\substack{j \in \mathcal{N}_{tv} \\ j \neq i}} x_{ij}^{tv} = 1 \qquad \forall i \in \mathcal{N}_{tv} \tag{44}$$

$$\sum_{\substack{i \in \mathcal{N}_{tv} \\ i \neq j}} x_{ij}^{tv} = 1 \qquad \forall j \in \mathcal{N}_{tv} \tag{45}$$

$$\sum_{i \in \mathcal{S}} \sum_{\substack{j \in \mathcal{S} \\ j \neq i}} x_{ij}^{tv} \leq |\mathcal{S}| - 1 \qquad \forall\, \mathcal{S} \subseteq \mathcal{N}_{tv} \setminus \{0\},\ |\mathcal{S}| \geq 2 \tag{46}$$

$$x_{ij}^{tv} \in \{0, 1\} \qquad \forall i, j \in \mathcal{N}_{tv},\ i \neq j. \tag{47}$$

Denote the optimal objective value of (43)–(47) by Z_{tv}^{det}. The master assignment $\mathcal{A}_{tv}$ is declared *time-feasible* if

$$Z_{tv}^{\text{det}} \leq T^{max}.$$

In our implementation, the subtour-elimination constraints (46) are treated as *lazy constraints*: they are not added a priori for all subsets, but are dynamically separated and added only when a violated subtour is detected.

Stochastic sub-problem (chance-constrained routing). In the stochastic setting, the routing feasibility is evaluated using the chance-constrained time budget (Definition 3), which yields conic quadratic constraints. Consequently, the sub-problem becomes a MISOCP. Since dedicated TSP solvers are not applicable in this setting, we rely on a commercial solver that natively supports conic constraints. As in the deterministic case, subtours are handled via lazy subtour-elimination constraints, added on demand during the solver's branch-and-cut.

Sub-problem solution strategy and parallelization. For the deterministic sub-problems, we leverage three solution engines: (i) a commercial solver (Gurobi) for the MIP/TSP model, (ii) the Concorde TSP solver (Applegate et al., 2006) (exact), and (iii) the Lin–Kernighan–Helsgaun (LKH) heuristic (Helsgaun, 2015). We first run LKH due to its speed; if the resulting tour has duration $\leq T^{max}$, we immediately certify feasibility. If LKH does not certify feasibility, we solve the minimum-duration TSP to optimality using Concorde to determine whether $Z_{tv}^{\det} \leq T^{max}$. A commercial solver remains available when additional modeling flexibility is required.

Finally, sub-problems are independent across (t, v). We therefore solve the $|\mathcal{T}| \times |\mathcal{V}$ routing sub-problems in parallel using multi-core execution, which substantially reduces wall-clock time in each LBBD iteration and accelerates the generation of feasibility cuts when infeasible assignments occur.

4.1.6 Algorithmic Framework

The proposed LBBD algorithm for the TRSP-STSU proceeds as follows.

Algorithm 1 LBBD for TRSP-STSU

Initialize: Apply uniform preprocessing (P1–P2; Section 4.1.4) and add the valid inequalities to the master.
Set optimality tolerance ϵ and maximum runtime $T_{\max}^{CPU}$.
Set incumbent (best feasible) value $Z^{LB} \leftarrow -\infty$.
while elapsed time $< T_{\max}^{CPU}$ and relative gap $> \epsilon$ **do**
    Explore the master problem within a branch-and-bound framework.
    Whenever an integer-feasible master solution $\bar{w}$ is found at a branch-and-bound node:
        let its objective value be Z^{UB};
        for each $(t, v) \in \mathcal{T} \times \mathcal{V}$ **in parallel**:

- Build the routing sub-problem on $\mathcal{A}_{tv} := \{i \in \mathcal{I} : \bar{w}_{itv} = 1\}$.
- **Deterministic:** solve the minimum-duration TSP and certify feasibility iff the optimal tour duration $\leq T^{max}$.
- **Stochastic:** solve the chance-constrained routing sub-problem (MISOCP) and certify feasibility accordingly.
- If infeasible: generate and add feasibility cuts (no-good / conflict-set cuts) as *lazy constraints*.

    If all sub-problems are feasible:

- Recover a complete solution (w, y, x) and compute its objective value Z.
- Update incumbent: $Z^{LB} \leftarrow \max\{Z^{LB}, Z\}$.

    Update relative gap:

$$\text{gap} \; = \; \frac{Z^{UB} - Z^{LB}}{|Z^{UB}| + \epsilon}.$$

end while
Output: Best feasible assignment and routing plan (w^*, y^*, x^*) with gap $\leq \epsilon$, or the best solution found within $T_{\max}^{CPU}$.

This iterative scheme alternates between a strengthened master assignment problem and independent routing sub-problems. Infeasible assignments are eliminated via dynamically generated feasibility cuts, while parallel sub-problem evaluation across (t, v) reduces the wall-clock time per iteration and accelerates convergence.

4.2 Branch-and-Cut (BC)

The BC algorithm serves as a rigorous baseline for our computational experiments. We employ the BC algorithm to solve the proposed models, iteratively incorporating subtour elimination constraints and other problem-specific cuts throughout the branch-and-bound process. Initially, subtour elimination constraints are added to the linear programming (LP) relaxation of the problem and are dynamically generated at subsequent nodes in the search tree. Whenever a solution at any node violates subtour elimination constraints, we generate the corresponding constraints and add them to the model, after which the problem is re-optimized at that node. This iterative procedure ensures that all solutions satisfy the subtour elimination constraints, progressively refining the feasible region of the problem. To efficiently detect and manage these violations, we employ the separation procedure proposed by Lysgaard et al. (2004), which is specifically tailored for the VRP. This enables the BC algorithm to maintain a feasible and optimized solution structure throughout the solution process. In addition to the model-agnostic inequalities of Section 4.1.3, we use exactly one extra preprocessing *only* in the monolithic BC/MIP (and not in the LBBD master nor in its TSP subproblems):

(P3) Time-infeasible two-customer arcs. We assume that travel times satisfy the triangle inequality and that all service times are nonnegative. Under these mild and standard conditions, any feasible route that visits customer i before j must take at least the time required for the simple depot–i–j–depot tour. Therefore, if this shortest possible two-customer tour already exceeds the time budget, the directed arc (i, j) can never appear in a feasible route:

$$\bar{\tau}_{0i} + \bar{S}_i + \bar{\tau}_{ij} + \bar{S}_j + \bar{\tau}_{j0} > T^{max} \implies x_{ijtv} = 0 \quad \forall\, t \in \mathcal{T},\, v \in \mathcal{V}.$$

This pruning rule safely eliminates time-infeasible arcs without excluding any route that could satisfy the duration constraint.

5 Numerical Results

In this section, we present the computational experiments conducted to evaluate the effectiveness of the proposed scheduling and routing models and solution approaches described in Section 4. The objectives are to assess the performance of the deterministic and stochastic models under various settings, analyze the impact of uncertainty and diminishing returns over time, and derive managerial insights that can guide decision-making in real-world operations. The experiments were run on a 64-bit Windows Server with two 2.4 GHz Intel Xeon CPUs and 24 GB RAM. The algorithms are implemented using the Python programming language and the Gurobi Solver version 11.0.3.

We generated a comprehensive set of instances using customer locations derived from real data provided by our industrial partner, Hydro-Québec. The test bed consists of four instance sets that feature 10, 20, 50, and 100 customers, respectively. Customers are categorized into prioritized and optional sets, maintaining a ratio of 3:2. Technicians in each instance are assigned heterogeneous skill levels across five skill types with integer levels ranging from 1 to 5. To reflect the hybrid skill-feasibility structure used in our models, we partition the skill set as $\mathcal{M} = \mathcal{M}^{\text{add}} \cup \mathcal{M}^{\text{cov}}$, where $\mathcal{M}^{\text{add}}$ contains additive (team-capacity) skills and $\mathcal{M}^{\text{cov}}$ contains qualification (coverage) skills that must be held by at least one crew member. In all instances, we set $|\mathcal{M}^{\text{add}}| = 3$ and $|\mathcal{M}^{\text{cov}}| = 2$. Customer skill requirements are generated accordingly: for $m \in \mathcal{M}^{\text{add}}$, the required level R'_{im} is interpreted as a team-capacity threshold and is sampled so that it is compatible with the maximum skill deliverable by a size-C team (for our experiments with $C = 2$, this corresponds to a maximum pooled level of 10 per skill). For $m \in \mathcal{M}^{\text{cov}}$, the requirement R'_{im} is interpreted as an individual qualification threshold, and we precompute the indicator $Q_{ilm} = 1$ if $R_{lm} \geq R'_{im}$ (and 0 otherwise) for use in the hybrid coverage constraints. Vehicles are integrated into the datasets based on customer volume, with fleet sizes set to 2 for 10 customers; 2 and 3 for 20 customers; 8 and 10 for 50 customers; and 10 and 15 for 100 customers. Each vehicle has the capacity to transport a maximum of two technicians, with $C = 2$ as a uniform constraint.

The time horizon is discretized into 2, 3, or 5 daily periods. Travel times between customer locations are derived from distances computed using customer coordinates and are subsequently rescaled to reflect realistic travel durations ranging from 10 to 60 minutes. Customer service times are assigned within a range of 30 to 120 minutes according to historical data, capturing differences in task complexity.

For reproducibility, customer benefit values G_{it} are generated following the structure provided by our industrial partner, Hydro-Québec. Each customer i is first assigned a base benefit G_{i1} sampled uniformly from the interval $[50, 500]$, which reflects the benefit ranges and relative priorities observed in historical operations. Diminishing returns across periods are then enforced through a customer-specific decay factor $\rho_i \sim \mathcal{U}[0.7, 1]$ drawn once for each customer and applied across the planning horizon, yielding

$$G_{it} = G_{i1}\,\rho_i^{t-1}, \qquad t = 1, \ldots, |\mathcal{T}|,$$

so that G_{it} is non-increasing over time for every customer. This formulation allows different customers to exhibit heterogeneous decay patterns while preserving the monotonic structure of service benefits across periods. The same random seed is used during instance generation to ensure reproducibility across deterministic, baseline stochastic, and high-uncertainty scenarios, so that differences between scenarios arise solely from the uncertainty parameters rather than underlying instance structure.

To incorporate real-world uncertainty, we introduce variability in both travel and service times, modeled as independent, normally distributed random variables. The mean values for these distributions are derived from deterministic travel and service times, with standard deviations ranging between 1% and 10% of the mean values. It must be noted that within each instance, travel/service time means and variances are constant across periods (stationary across t). For all experiments with the stochastic model, chance constraints are enforced at confidence level $\alpha = 0.95$. In the baseline stochastic scenario, the standard deviations for travel and service times are sampled uniformly between 1% and 10% of the mean values. In the high-uncertainty scenario, the range is increased to 10%–20%, providing a stricter stress test of robustness. To ensure diversity in the dataset and robustness in performance evaluation, we created multiple instances per configuration. For each combination of customer size, time horizon (2, 3, or 5 days), and fleet size, we generated 10 different versions that vary in customer demand, technician pool characteristics (size and skill profiles), and travel conditions, while preserving the core structure of the instance. This setup results in a total of 210 instances, as summarized in Table 2.

Prior to solving, we apply the uniform preprocessing to every instance, including the skill-based filters implied by the hybrid feasibility structure (additive-capacity and qualification coverage) and the time-feasibility checks. The same reduced instance is used for both algorithms (BC and LBBD) to ensure methodological parity.

Table 2: Summary of instance generation

Number of Customers	Fleet Size	Time Horizon (days)	Instances per Setting	# Instances
10	2	2, 3, 5	10	$1 \times 3 \times 10 = 30$
20	2, 3	2, 3, 5	10	$2 \times 3 \times 10 = 60$
50	8, 10	2, 3, 5	10	$2 \times 3 \times 10 = 60$
100	10, 15	2, 3, 5	10	$2 \times 3 \times 10 = 60$
Total				**210**

5.1 Performance of the Algorithms

This subsection presents a comparative performance analysis of the Logic-Based Benders Decomposition (LBBD) and Branch-and-Cut (BC) approaches applied to the 210 benchmark instances summarized in Table 2. The algorithms are evaluated using three key performance indicators: the number of instances solved to proven optimality, the average optimality gap at termination, and the average computational time. All algorithms are executed under identical computational settings with a uniform time limit of 3600 seconds per instance, ensuring a fair and consistent comparison.

Altogether, the LBBD approach demonstrates superior performance among the exact methods in terms of solution quality, robustness, and scalability. Across the full test set, LBBD solves 202 out of 210 instances to optimality, corresponding to a very high success rate. The overall average optimality gap is limited to 6.8%, indicating that even when optimality is not proven, the algorithm produces near-optimal solutions. A more detailed examination across instance sizes shows that LBBD solves all small instances (10 and 20 customers) to optimality, achieving a zero average gap in these groups. For the 50-customer instances, LBBD solves 58 out of 60 cases optimally with an average gap of 9.2%. For the largest instances with 100 customers, the algorithm still solves 54 instances to optimality, with the average gap increasing to 14.5%. This gradual increase reflects the expected growth in combinatorial complexity while still indicating strong scalability of the decomposition strategy.

In contrast, the BC approach exhibits weaker performance overall. Across all instances, BC solves 161 cases to optimality and records a substantially larger average optimality gap of 31.8%. The difference between the methods becomes particularly pronounced as instance size increases. In the 20-customer group, BC solves only 31 out of 60 instances optimally and leaves an average gap of 23.2%. For 50-customer instances, although BC solves 55 instances to optimality, the remaining cases produce a large average gap of 34.4%. The deterioration is most significant for the 100-customer group, where the average gap rises to 51.2%, indicating difficulty in closing the bound within the allowed time.

From a computational efficiency perspective, LBBD also provides a favorable runtime profile. The weighted average runtime across all 210 instances is 1,616 seconds. Runtime increases with instance size, as expected, but remains controlled. LBBD requires on average only 65 seconds for 10-customer instances and 950 seconds for 20-customer instances. For 50-customer problems, the runtime increases to 2,220 seconds, and for the largest instances the average runtime reaches 2,452 seconds. These results suggest that the decomposition effectively reduces search complexity by separating high-level assignment decisions from detailed feasibility checks.

The BC approach requires significantly more computational effort overall, with an average runtime of 2,493 seconds. Even for small instances, BC is slower than LBBD, requiring 295 seconds on average for 10 customers. The gap widens for 20-customer instances, where BC nearly reaches the time limit on average (3,150 seconds). Although BC shows a slightly lower runtime than LBBD for the 100-customer group, this is primarily due to earlier termination with large remaining optimality gaps rather than improved computational efficiency.

Table 3 summarizes the performance of both algorithms across all instance groups.

The results reveal several consistent trends. First, both approaches perform well on small instances, although LBBD achieves optimal solutions faster. Second, as instance size

Table 3: Comparison of LBBD and BC across all instance groups (3600s time limit)

Number of Customers	# Instances	Algorithm	# Opt. Solved	Avg. Gap (%)	Avg. Time (s)
10	30	**LBBD**	**30**	**0.0**	**65**
		BC	25	4.8	295
20	60	**LBBD**	**60**	**0.0**	**950**
		BC	31	23.2	3,150
50	60	**LBBD**	**58**	**9.2**	**2,220**
		BC	55	34.4	3,147
100	60	**LBBD**	**54**	**14.5**	**2,452**
		BC	50	51.2	2,281
Total / Average		**LBBD**	**202**	**6.8**	**1,616**
		BC	161	31.8	2,493

increases, the advantage of the decomposition approach becomes more pronounced: LBBD maintains relatively low optimality gaps and a high optimality rate, while BC experiences a significant deterioration in gap. Third, runtime patterns indicate that LBBD achieves a more effective balance between exploration and convergence, whereas BC often spends substantial computational effort strengthening the search tree without fully closing the optimality gap.

Altogether, LBBD exhibits a pattern of graceful degradation as problem size increases: optimality rates decrease moderately, optimality gaps grow gradually, and runtimes remain within practical limits. In contrast, BC shows weaker scalability and consistently larger gaps. These findings highlight the effectiveness of exploiting problem structure through logic-based decomposition and confirm that LBBD provides a more robust and scalable exact solution framework for larger and more complex instances.

5.2 Scenario Design and Sensitivity Analysis Framework

Given that LBBD demonstrated the best overall performance in terms of solution quality, computational efficiency, and robustness across all 210 instances, it is selected as the solution method for the scenario and sensitivity analysis. Unless otherwise specified, all results are evaluated relative to the *baseline model*, which corresponds to the same data environment used in the algorithmic comparison in Section 5.1. The scenarios are designed to assess the sensitivity of planning outcomes to key operational dimensions, including elevated uncertainty, improvements in technician skill levels (+1 and +3), and restricted maximum working hours of seven per day instead of eight. Together, these configurations reflect common sources of variability and managerial levers in technician scheduling.

System outcomes are evaluated using complementary metrics that capture economic, service, and resource-allocation effects at the planning stage. The first metric, *Total Benefit*, measures the economic outcome of a solution and is computed as the service benefits of completed visits minus negligible operating costs; higher values indicate that the model allocates resources to maximize economic returns. The second metric, *Optional Visit Coverage*, quantifies the proportion of optional customers that are scheduled for service within the planning horizon. It is formally defined as

$$\text{Optional Visit Coverage} = \frac{\sum_{i \in \mathcal{I}^o} \sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}} w_{itv}}{I^o} \times 100, \tag{48}$$

where $w_{itv} = 1$ if optional customer i is visited in period t by vehicle v. By construction, all prioritized customers $\mathcal{I}^p$ are visited exactly once; thus, Optional Visit Coverage is evaluated only over the optional set. The third metric, *Technician Utilization*, captures workforce efficiency and is defined as the average proportion of available working hours assigned to tasks; high utilization signals dense scheduling, while lower values indicate slack introduced by constraints or uncertainty buffers.

5.2.1 Impact of Uncertainty on Planning Outcomes

Uncertainty in travel and service times plays a decisive role in shaping system performance. To assess its impact, we analyze three configurations side by side: (i) a deterministic benchmark with fixed times, (ii) the baseline stochastic model reflecting moderate variability, and (iii) a high-uncertainty case with variability increased to 10–20%. The aggregated outcomes over 210 instances are reported in Table 4.

The deterministic model, free of variability, provides the natural upper bound for nominal performance. Under predictable conditions, the system exploits available capacity more aggressively, yielding an average total benefit of 13,610 units. Technician utilization reaches 87.0%, while Optional Visit Coverage averages 85.6%, indicating that the planner prioritizes higher-value tasks rather than maximizing the breadth of optional service. Computational effort remains comparatively lower and more stable than in the stochastic settings, with a mean solution time of 905 seconds. These results confirm that predictable operating conditions enable tighter schedules and higher nominal profitability.

Introducing stochasticity through the baseline model alters planning behavior substantially. The average benefit declines to 12,304 units, reflecting the inclusion of safety buffers that hedge against variability in travel and service times. This reduction should be interpreted as a deliberate trade-off rather than a loss of model quality. In contrast to the deterministic formulation, Optional Visit Coverage increases markedly to 95.6% on average. This shift suggests that the uncertainty-aware planner expands the set of optional visits by favoring tasks with more predictable durations that can be accommodated within conservative schedules.

Technician utilization, however, decreases to 71.0%, revealing the intentional slack embedded in the schedules to preserve feasibility under uncertainty. The combination of higher optional coverage and lower utilization indicates a structural change in task selection: the model balances economic value with temporal reliability, selecting optional activities that fit safely within buffered schedules while leaving additional workforce capacity unused to protect against time overruns. At the same time, computational effort increases relative to the deterministic case, with mean solution times rising to approximately 1,616 seconds, reflecting the added complexity introduced by uncertainty modeling.

The high-uncertainty scenario (10–20%) amplifies these effects. Average benefit declines further to 10,597 units, representing a reduction of roughly 22% relative to deterministic conditions and about 14% compared to the baseline stochastic model. Optional Visit Coverage decreases to 89.5%, indicating that stronger variability limits the set of optional tasks that can be scheduled without compromising feasibility. This behavior reflects a robustness-oriented strategy in which the model becomes increasingly selective as uncertainty grows, prioritizing temporal stability over service breadth.

The most pronounced impact of rising uncertainty is observed in technician utilization, which falls to 57.1% under high variability. This decline illustrates how additional uncertainty forces the algorithm to leave a substantial portion of available capacity unused rather than risking infeasibility. Computational effort also increases significantly: mean solution time grows to 1,822 seconds and runtime variability expands, highlighting the algorithmic difficulty of maintaining feasibility guarantees under large variability.

Across scenarios, a consistent pattern emerges. Profitability decreases as uncertainty intensifies, reflecting the cost of adopting more conservative and robustness-oriented schedules. Resource utilization exhibits a monotonic decline from deterministic to baseline and high-uncertainty settings, confirming that uncertainty primarily affects how tightly workforce capacity can be scheduled. Optional Visit Coverage follows a non-monotonic pattern: moderate uncertainty encourages broader inclusion of optional tasks, while high variability reverses this effect and leads to more selective planning. This non-monotonic behavior does not contradict the deterministic benchmark once the objective structure is considered. Because the model maximizes total benefit rather than optional coverage, deterministic planning may concentrate capacity on a smaller set of high-value optional activities instead of maximizing the number of visits. Under chance constraints, each optional task consumes not only its mean duration but also a variability-driven safety margin, effectively increasing its time cost. As a result, time-intensive or highly variable optional tasks become less attractive, and the stochastic planner may replace a few large activities with multiple shorter and more temporally reliable ones. This substitution mechanism explains how optional coverage can increase under moderate uncertainty even as total benefit declines, while stronger variability eventually restricts both benefit and coverage as safety buffers dominate available capacity.

From a computational perspective, stochasticity introduces additional complexity but remains manageable for the LBBD framework. Solution times increase with uncertainty and exhibit greater dispersion, yet the algorithm continues to solve all instances within practical limits. Altogether, the results demonstrate that the LBBD approach adapts to increasing uncertainty by progressively embedding slack and prioritizing feasibility and robustness over nominal utilization and benefit, highlighting the expected trade-off between efficiency and reliability in uncertainty-aware planning.

Table 4: Comparative planning outcomes under different uncertainty levels (210 instances)

Scenario	Metric	Min	Max	Mean	Std. Dev.
Deterministic	Total Benefit	3,243	33,078	13,610	7,458
	Optional Visit Coverage (%)	56.9	100.0	85.6	5.8
	Technician Utilization (%)	76.6	100.0	87.0	7.9
	Computational Time (s)	60.10	3,605.98	905.12	901.47
Baseline Model	Total Benefit	2,060	31,111	12,304	7,262
	Optional Visit Coverage (%)	67.9	100.0	95.6	3.7
	Technician Utilization (%)	53.2	86.2	71.0	8.2
	Computational Time (s)	70.15	3,611.93	1,616.41	902.95
High Uncertainty (10–20%)	Total Benefit	1,848	28,150	10,597	6,575
	Optional Visit Coverage (%)	76.9	100.0	89.5	7.8
	Technician Utilization (%)	46.1	73.1	57.1	6.8
	Computational Time (s)	56.18	3,630.82	1,821.63	1,407.66

5.2.2 Effect of Technician Skill Levels

This analysis investigates the influence of technician skill enhancement on overall system performance. Two skill uplift configurations are considered: (i) increasing all technician skill levels by one point, and (ii) increasing them by three points. These adjustments reflect real-world interventions such as training programs, certifications, or accumulated experience that expand technician capability and task eligibility.

The results, summarized in Table 5, indicate that improving technician skills produces consistent but nuanced operational gains. A one-level skill increase raises average total benefit from 12,304 units in the baseline to 13,237 units, while a three-level increase further improves benefit to 13,723 units. The magnitude of improvement is moderate, suggesting that the baseline system already captures a substantial portion of feasible task–technician matches and that additional skill flexibility primarily refines scheduling efficiency rather than fundamentally expanding feasible assignments.

Optional Visit Coverage increases sharply with skill uplift. In the baseline, average coverage is 95.6%, indicating that a nontrivial fraction of optional demand remains unserved under existing skill constraints. With either a +1 or +3 increase, Optional Visit Coverage rises to 99.9%, demonstrating that enhanced skills remove binding eligibility bottlenecks and allow the model to accommodate nearly all optional visits within the planning horizon. The absence of further improvement between the +1 and +3 cases indicates diminishing returns: most coverage gains are captured through modest training improvements.

The most pronounced effect of skill enhancement is observed in technician utilization. Utilization rises from 71.0% in the baseline to 86.8% with a +1 skill increase and stabilizes at 87.2% under the +3 configuration. This pattern confirms that skill constraints constitute a major limiting factor in workforce deployment: once technicians become capable of performing a broader set of tasks, schedules can be tightened and idle capacity reduced. The limited additional gain between the +1 and +3 scenarios further reinforces the presence of diminishing marginal returns to broad upskilling.

Computational performance remains within practical limits across all scenarios. The +1 configuration slightly decreases average runtime to 1,423 seconds, reflecting the expanded assignment flexibility that must be explored during optimization. Interestingly, the +3 configuration reduces runtime to 1,003 seconds on average, suggesting that once skill constraints become largely nonbinding, the solver benefits from a smoother feasible region and faster convergence. In summary, the findings highlight technician skill development as an effective managerial lever: even modest skill improvements can substantially increase optional service coverage and workforce utilization, while delivering moderate gains in total benefit.

Table 5: Performance under Increased Technician Skill Levels

Skill Level Increase	**Total Benefit**	**Optional Visit Coverage (%)**	**Technician Utilization (%)**	**Time (s)**
Baseline	12,304	95.6	71.0	1,616
+1	13,237	99.9	86.8	1,423
+3	13,723	99.9	87.2	1,003

5.2.3 Impact of Maximum Work Hours

In this scenario, the maximum allowable working time per technician per day is reduced from 8 hours to 7 hours. This modification simulates regulatory, contractual, or ergonomic constraints that limit daily shift durations. The objective is to evaluate how tighter labor regulations influence operational performance and resource allocation within the scheduling framework.

As reported in Table 6, reducing working hours leads to a decline in average total benefit from 12,304 to 11,614 units. This reduction reflects the smaller feasible service set that can be accommodated within the compressed time budget. With less available time per technician, the planner must prioritize higher-value and more time-efficient assignments, inevitably leaving some optional opportunities unserved. Optional Visit Coverage decreases from 95.6% to 92.1%, indicating that the tighter time horizon primarily affects discretionary service capacity while prioritized customers remain fully satisfied by design. The magnitude of this drop highlights how sensitive optional service provision is to working-time restrictions, even when baseline coverage is already high.

Technician utilization increases from 71.0% to 74.3% under reduced hours. This increase reflects denser schedules within the shorter working window: although the absolute amount of performed work decreases, the available time is used more intensively. In other words, the constraint reduces idle capacity but simultaneously limits the total volume of service that can be delivered. Computational effort decreases moderately, with average runtime decreasing from 1,616 to 1,607 seconds. The additional shift constraint tightens the feasible region and requires the solver to explore more complex trade-offs between task selection and temporal feasibility, but overall tractability remains within practical limits.

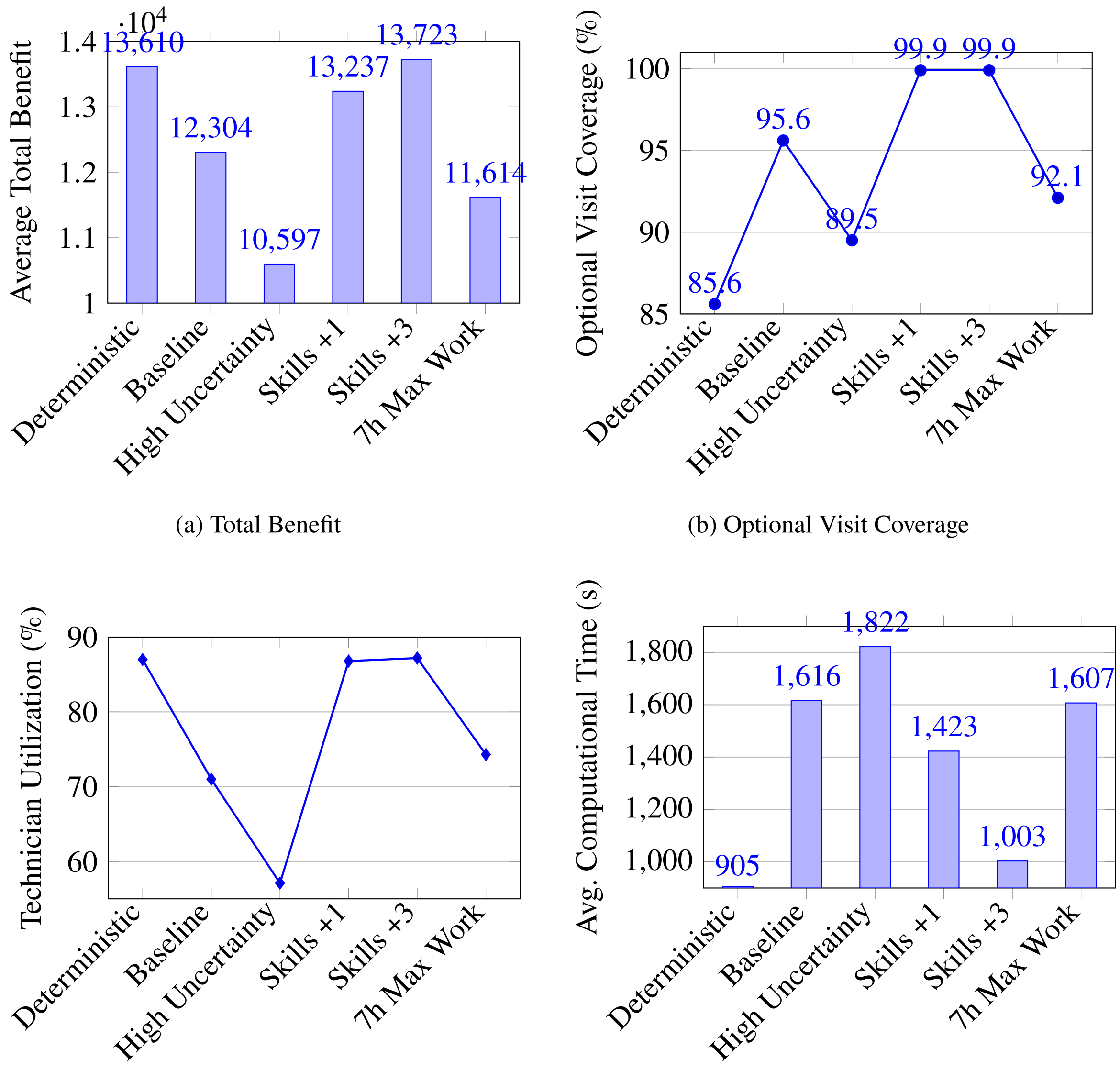
In summary, the findings indicate that reducing maximum work hours imposes a clear trade-off between labor regulation and operational efficiency. While feasibility is preserved and workforce time is used more intensively, both total benefit and optional service capacity decline. These results suggest that policies limiting daily working time should be accompanied by compensatory measures—such as increased staffing, improved skills, or process efficiencies—to mitigate their impact on service performance.

Table 6: Performance under Reduced Maximum Work Hours

Metric	**Baseline (8h)**	**Reduced (7h)**
Total Benefit	12,304	11,614
Optional Visit Coverage (%)	95.6	92.1
Technician Utilization (%)	71.0	74.3
Computational Time (s)	1,616	1,607

5.3 Visual Insights from Scenario Results

This section complements the tabular results by providing comparative visualizations that summarize the effects of different modeling assumptions and operational parameters. The figures present aggregated outcomes over all 210 benchmark instances and highlight systematic trends in total benefit, Optional Visit Coverage, technician utilization, and computational effort across scenarios.



(a) Total Benefit

(b) Optional Visit Coverage

(c) Technician Utilization

(d) Computational Time

Figure 1: Scenario comparisons for four planning-stage KPIs.

Figure 1a presents the average total benefit across scenarios. The deterministic setting yields the highest nominal benefit among uncertainty settings, while the baseline stochastic model produces a moderate reduction due to the inclusion of robustness buffers. Increasing technician skills improves benefit relative to the baseline, with *Skills +3* providing the largest gain, though with diminishing returns compared to *Skills +1*. In contrast, *High Uncertainty* and the *7h Max Work* scenario both reduce profitability, with elevated uncertainty producing the strongest decline.

Optional Visit Coverage is summarized in Figure 1b. The results reveal a non-monotonic pattern across scenarios. Coverage increases from 85.6% in the deterministic setting to 95.6% under baseline uncertainty, suggesting that moderate variability encourages the inclusion of optional tasks that can be scheduled safely within buffered plans. Under high uncertainty, coverage declines to 89.5%, indicating that stronger variability restricts the set of optional visits that can be accommodated without compromising feasibility. Skill-enhanced scenarios achieve near-complete coverage (99.9%), demonstrating that eligibility constraints become the dominant bottleneck once technician capabilities expand. Reduced working hours lead to a moderate decline to 92.1%, highlighting the sensitivity of optional service provision to available workforce time.

Figure 1c highlights workforce usage patterns. Utilization is highest in the deterministic and skill-enhanced scenarios (around 87%), demonstrating that relaxed skill constraints enable tighter schedules comparable to fully predictable conditions. Baseline stochastic planning introduces slack, reducing utilization to 71.0%, while high uncertainty further lowers it to 57.1%. Shortened working hours increase utilization slightly relative to the baseline, as tasks must be packed more densely within the reduced horizon.

Average computational time is presented in Figure 1d. As expected, higher uncertainty increases runtime due to the added complexity of maintaining feasibility under variability. Skill-enhanced scenarios exhibit mixed computational effects: the +1 case slightly decreases runtime, whereas the +3 configuration reduces it, suggesting that largely non-binding skill constraints facilitate faster convergence. Altogether, runtimes remain within practical limits across all configurations, confirming the scalability and robustness of the LBBD framework under diverse operational conditions.

5.4 Case Study: Hydro-Québec Application

To complement the benchmark experiments and scenario analyses presented in the previous sections, we now examine a large-scale case study based on data provided by our industrial partner, Hydro-Québec. The purpose of this case study is twofold. First, it illustrates the practical applicability of the proposed modeling framework and LBBD solution approach in a realistic industrial setting. Second, it evaluates how deterministic and chance-constrained planning decisions perform when subjected to adverse execution conditions through out-of-sample simulation.

The case study considers 200 geographically distributed customers to be served over a multi-period planning horizon. Customers are divided into prioritized and optional sets, maintaining the same 3:2 ratio used throughout the numerical experiments. The prioritized customers must be visited exactly once within the horizon, while optional customers are served only if sufficient capacity is available. The system operates with 15 vehicles, each staffed by exactly two technicians per period for three periods, resulting in coordinated mobile teams. Technicians possess heterogeneous skill profiles across five skill types, which constrain task eligibility in the same manner as in the benchmark instances. Customer locations are taken from the partner dataset, and travel times are computed from inter-location distances and scaled to realistic durations (10–60 minutes). Service times are drawn from the corresponding service-time field in the dataset and reflect task-dependent workload requirements.

In the stochastic setting, both travel and service times are modeled as independent normally distributed random variables with mean values equal to their deterministic counterparts. The chance-constrained formulation enforces feasibility at confidence level $\alpha = 0.95$. Two planning policies are evaluated. The *deterministic policy* assumes that travel and service times are known with certainty and is implemented by setting all standard deviations to zero, optimizing against nominal mean values. The *chance-constrained policy* explicitly accounts for uncertainty by embedding safety margins in routing and scheduling decisions. Both policies are solved using the LBBD algorithm described in Section 4, with identical preprocessing, time limits, and algorithmic parameters. The objective in both cases is to maximize the total planned benefit subject to service, skill, and team-assignment constraints.

In the planning stage, both policies achieve nearly identical objective values. The deterministic policy yields a planned total benefit of 30,409.27, while the chance-constrained policy attains 30,353.84. This small difference indicates that the robustness adjustments introduced by the chance-constrained formulation have only a minor impact on nominal planning performance. Importantly, similar planning-stage objective values do not imply similar execution-stage behavior. As emphasized throughout this paper, deterministic metrics alone do not capture execution risk or the cost of schedule infeasibility under uncertainty. To evaluate these aspects, we next assess both plans under a controlled stress-test execution environment.

To evaluate robustness, we subject both fixed plans to an out-of-sample Monte Carlo simulation that mimics adverse execution conditions by imposing a tight working-time budget that induces systematic pressure on route feasibility. In the reported stress test, the per-route working-time limit used in simulation is $T_{\max} = 360$ minutes, yielding a stringent execution regime that amplifies the impact of stochastic travel and service times and makes schedule robustness critical. The reported simulation results correspond to the evaluated daily execution slice of the case-study plan. To allow vehicles to remain idle while preserving model feasibility, we adopt the dummy-customer extension described in Remark 1. Vehicles assigned only to dummy visits are not counted as active slots; thus, the reported number of active slots reflects vehicles performing at least one real customer visit. For each planning policy, we simulate 1,000 independent execution days. In each simulation run, travel and service times are independently sampled from their stochastic distributions, and the realized duration of each planned route is computed. The *Realized Benefit* is defined as the planned benefit minus penalties incurred whenever a route exceeds the allowable working-time limit. Penalty costs are assessed proportionally to overtime duration. In addition to average performance, we report summary indicators describing route-level reliability, overtime severity, workload dispersion, and resource utilization.

Table 7 reports the execution-stage performance of both planning policies under this stringent stress test. Despite their nearly identical planned objectives, the two policies exhibit markedly different realized outcomes. The deterministic plan achieves an average realized benefit of 25,228.37, whereas the chance-constrained plan improves this value to 26,706.70. Consistently, the average daily penalty decreases from 5,180.90 under the deterministic plan to 3,647.14 under the chance-constrained plan, confirming that explicitly accounting for uncertainty during planning yields materially better realized performance. Route-level reliability highlights this contrast. Under the deterministic plan, 66.60% of routes exceed the time limit, compared with 42.86% under the chance-constrained policy. Average overtime decreases from 57.57 minutes to 34.73 minutes, and the upper-tail overtime indicator also improves, with $\mathrm{VaR}_{0.95}$ decreasing from 135.16 to 121.14 minutes. The same pattern appears in the frequency of more severe violations: the critical overtime rate, defined as the percentage of routes exceeding the limit by more than 15 minutes, falls from 66.42% to 42.74%. Minor overtime events remain negligible under both policies, with rates of only 0.03% and 0.01%, respectively.

To further characterize workload dispersion, Table 7 reports the coefficient of variation of realized route durations. The deterministic plan yields a Route Duration CV (standard

Table 7: Hydro-Québec case study under stress-test execution.

Metric	**Deterministic Plan**	**Chance-Constrained Plan**
Planned Benefit	30,409.27	30,353.84
Realized Benefit (avg.)	25,228.37	26,706.70
Penalty (avg./day)	5,180.90	3,647.14
Total Duration Violation (%)	66.60	42.86
Avg. Overtime (min)	57.57	34.73
Overtime $VaR_{0.95}$ (min)	135.16	121.14
Critical Overtime Rate (> 15 min) (%)	66.42	42.74
Minor Overtime Rate (< 5 min) (%)	0.03	0.01
Route Duration CV	0.3482	0.3852
Active Slots	6	7
Fleet Utilization (%)	40.04	40.79
Active Utilization (%)	100.00	87.41

Notes: Route Duration CV denotes the coefficient of variation (standard deviation divided by mean) of realized route durations across simulated routes. Fleet Utilization is reported relative to the total available fleet time in the evaluated day (15×360 minutes), whereas Active Utilization is reported relative only to the time capacity of active slots.

deviation divided by mean) of 0.3482, whereas the chance-constrained plan has a slightly higher value of 0.3852. This indicates somewhat greater relative dispersion in realized route durations under the stochastic-aware plan. Nevertheless, this modest increase in route duration variability is accompanied by substantially lower overtime frequency and severity, suggesting that the chance-constrained policy accepts a slightly more uneven distribution of workload in exchange for materially improved protection against route overruns.

The utilization indicators provide additional operational insight. The deterministic plan uses 6 active slots, while the chance-constrained plan uses 7. Correspondingly, fleet utilization increases slightly from 40.04% to 40.79%. At the same time, active utilization decreases from 100.00% to 87.41%, indicating that the chance-constrained plan spreads work over more active vehicle-period slots and preserves more slack within each active route. This slack appears to play an important role in absorbing stochastic travel and service-time variability during execution.

In total, the case study highlights the limitations of evaluating plans solely on deterministic or planning-stage metrics. Although the deterministic policy appears competitive when judged only by nominal benefit, it produces schedules that incur substantially larger penalties and more frequent overtime violations once uncertainty is realized. By contrast, the chance-constrained policy internalizes execution risk during planning and yields schedules that are more stable in a demanding operational environment. From a managerial perspective, these results emphasize that robustness can deliver meaningful gains even when planning-stage objectives are nearly identical. Reductions in overtime frequency and severity translate directly into improved cost control, greater schedule predictability, and more reliable service delivery. The Hydro-Québec case thus demonstrates that incorporating uncertainty awareness into planning can improve realized performance without materially sacrificing nominal efficiency, particularly in technician-intensive service systems where execution deviations are difficult to correct in real time.

6 Conclusion

We studied a real-world technician routing and scheduling problem motivated by Hydro-Québec that integrates vehicle routing, multi-period team assignment, heterogeneous technician skills (with a hybrid additive–coverage feasibility structure), customer priorities, and stochastic travel and service times under chance-constrained feasibility requirements. To solve this problem at scale, we developed an integrated optimization framework and evaluated two exact solution strategies: a direct Branch-and-Cut (BC) approach and a tailored Logic-Based Benders Decomposition (LBBD) algorithm.

Across 210 benchmark instances spanning 10–100 customers, multiple fleet sizes, and 2–5 daily periods, LBBD consistently outperformed BC in both solution quality and efficiency. Under a 3600 s time limit, LBBD solved 202 instances to optimality (vs. 161 for BC) and achieved a substantially smaller average optimality gap (6.8% vs. 31.8%). LBBD also reduced mean runtime from 2,493 s (BC) to 1,616 s overall, and exhibited more graceful degradation as instance size increased (e.g., average gaps of 9.2% for 50-customer instances and 14.5% for 100-customer instances), highlighting its scalability and robustness.

Scenario and sensitivity analyses using LBBD provided actionable managerial insights. Comparing planning outcomes across uncertainty regimes revealed a clear efficiency–reliability trade-off. The deterministic benchmark achieved the highest nominal performance, with mean total benefit of 13,610 and technician utilization of 87.0%, alongside 85.6% Optional Visit Coverage. Introducing moderate uncertainty through the baseline chance-constrained model reduced mean total benefit to 12,304 and utilization to 71.0%, while increasing Optional Visit Coverage to 95.6%, indicating a shift toward time-reliable optional tasks and the deliberate inclusion of schedule slack. Under high uncertainty (10–20%), the system became more conservative: mean total benefit declined to 10,597, Optional Visit Coverage fell to 89.5%, and utilization dropped further to 57.1%, while average runtime increased to 1,822 s. Overall, uncertainty primarily tightens schedules through buffers that reduce utilization and profitability, and its effect on optional coverage is non-monotonic: moderate uncertainty broadens optional inclusion, whereas strong variability forces selectivity.

Skill enhancement emerged as a powerful lever for improving service breadth and workforce efficiency. Relative to the baseline (benefit 12,304; coverage 95.6%; utilization 71.0%), a uniform +1 skill uplift increased mean benefit to 13,237, raised Optional Visit Coverage to 99.9%, and increased utilization to 86.8%. A +3 uplift yielded further (but diminishing) gains, reaching a mean benefit of 13,723 with essentially identical coverage (99.9%) and utilization (87.2%). These results indicate that modest training investments can remove binding eligibility constraints and nearly eliminate unserved optional demand, while additional upskilling beyond that level yields smaller marginal improvements.

Restricting daily working hours from eight to seven produced a distinct trade-off between labor constraints and operational performance. Average total benefit decreased from 12,304 to 11,614 and Optional Visit Coverage declined from 95.6% to 92.1%, while utilization increased from 71.0% to 74.3% due to denser schedules within a shorter time budget. Runtime decreased moderately (from 1,616 s to 1,607 s), reflecting the tighter feasible region induced by the additional time restriction.

A large-scale Hydro-Québec case study with 200 customers and 15 vehicles over three periods further demonstrated the practical value of uncertainty-aware planning. While the deterministic and chance-constrained policies produced nearly identical planned benefits (30,409.27 vs. 30,353.84), their execution outcomes diverged substantially under a stringent stress-test simulation with $T_{\max} = 360$ minutes and 1,000 Monte Carlo replications. The chance-constrained plan improved average realized benefit from 25,228.37 to 26,706.70 and reduced average daily penalties from 5,180.90 to 3,647.14. Reliability improved markedly: the total duration violation dropped from 66.60% to 42.86%, and average overtime decreased from 57.57 to 34.73 minutes. These findings underscore that comparable planning-stage objective values can mask substantial differences in execution risk, and that incorporating chance constraints can yield more stable and cost-effective operations in practice.

In summary, this work advances technician routing and scheduling by unifying routing, multi-period team assignment, hybrid skill feasibility, and chance-constrained robustness within a decomposition-based optimization framework. The numerical results suggest that (i) explicitly modeling uncertainty is essential to manage execution risk, (ii) modest technician upskilling can dramatically improve optional service coverage and utilization, and (iii) tighter work-hour regulations reduce service capacity and profitability unless compensated by staffing or capability improvements. Future research directions include hybrid exact–heuristic accelerations for very large instances, real-time recourse and rescheduling under disruptions, and data-driven learning of travel/service-time uncertainty to better calibrate robustness levels.

Achnowlegment

The authors acknowledge the support of the Digital Research Alliance of Canada for providing the computational resources used for running the computational experiments. This work was supported by Mitacs through the Mitacs Accelerate program IT23679.

References

Applegate, D. L., R. E. Bixby, V. Chvatál, W. J. Cook. 2006. *The Traveling Salesman Problem: A Computational Study*. Princeton University Press.

Balas, E. 1975. Facets of the knapsack polytope. *Mathematical Programming* **8** 146–164. doi:10.1007/BF01580440.

Barrera, D., N. Velasco, C.-A. Amaya. 2012. A network-based approach to the multi-activity combined timetabling and crew scheduling problem: Workforce scheduling for public health policy implementation. *Computers & Industrial Engineering* **63**(4) 802–812.

Bostel, N., P. Dejax, P. Guez, F. Tricoire. 2008. Multiperiod planning and routing on a rolling horizon for field force optimization logistics. *The Vehicle Routing Problem: Latest Advances and New Challenges*. Springer, 503–525.

Bughin, J., M. Chui, J. Manyika. 2013. Ten IT-enabled business trends for the decade ahead. *McKinsey Quarterly* **13**(May) 1–13.

Cabrera, N., J.-F. Cordeau, J. E. Mendoza. 2026. A survey of workforce scheduling and routing problems. *Computers & Operations Research* 107344.

Castillo-Salazar, J. A., D. Landa-Silva, R. Qu. 2016. Workforce scheduling and routing problems: literature survey and computational study. *Annals of Operations Research* **239** 39–67.

Chen, L., M. Gendreau, M. H. Hà, A. Langevin. 2016a. A robust optimization approach for the road network daily maintenance routing problem with uncertain service time. *Transportation Research Part E: Logistics and Transportation Review* **85** 40–51.

Chen, X., M. Hewitt, B. W. Thomas. 2018. An approximate dynamic programming method for the multi-period technician scheduling problem with experience-based service times and stochastic customers. *International Journal of Production Economics* **196** 122–134.

Chen, X., K. Li, S. Lin, X. Ding. 2024. Technician routing and scheduling with employees' learning through implicit cross-training strategy. *International Journal of Production Economics* **271** 109208.

Chen, X., B. W. Thomas, M. Hewitt. 2016b. The technician routing problem with experience-based service times. *Omega* **61** 49–61.

Haugen, D. L., A. V. Hill. 1999. Scheduling to improve field service quality. *Decision Sciences* **30**(3) 783–804.

Helsgaun, K. 2015. Solving the equality generalized traveling salesman problem using the Lin–Kernighan–Helsgaun algorithm. *Mathematical Programming Computation* **7**(3) 269–287.

Jaillet, P., J. Qi, M. Sim. 2016. Routing optimization under uncertainty. *Operations Research* **64**(1) 186–200.

Khattara, A., W. R. Cherif-Khettaf, M. Mostefai. 2017. Variable neighborhood search procedures for the multi-period technician routing and scheduling problem. *2017 4th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 0839–0844.

Kovacs, A. A., S. N. Parragh, K. F. Doerner, R. F. Hartl. 2012. Adaptive large neighborhood search for service technician routing and scheduling problems. *Journal of Scheduling* **15** 579–600.

Li, H., S. Mi, Q. Li, X. Wen, D. Qiao, G. Luo. 2020. A scheduling optimization method for maintenance, repair and operations service resources of complex products. *Journal of Intelligent Manufacturing* **31** 1673–1691.

Lysgaard, J., A. N. Letchford, R. W. Eglese. 2004. A new branch-and-cut algorithm for the capacitated vehicle routing problem. *Mathematical Programming* **100** 423–445.

Mathlouthi, I., M. Gendreau, J.-Y. Potvin. 2021. A metaheuristic based on tabu search for solving a technician routing and scheduling problem. *Computers & Operations Research* **125** 105079.

Moradi, B. 2020. The new optimization algorithm for the vehicle routing problem with time windows using multi-objective discrete learnable evolution model. *Soft Computing* **24**(9) 6741–6769.

Nemhauser, G. L., L. A. Wolsey. 1988. *Integer and Combinatorial Optimization*. John Wiley & Sons.

Nowak, M., P. Szufel. 2024. Technician routing and scheduling for the sharing economy. *European Journal of Operational Research* **314**(1) 15–31.

Padberg, M., G. Rinaldi. 1991. A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. *SIAM Review* **33**(1) 60–100.

Paradiso, R., R. Roberti, D. Laganá, W. Dullaert. 2020. An exact solution framework for multitrip vehicle-routing problems with time windows. *Operations Research* **68**(1) 180–198.

Pekel, E. 2020. Solving technician routing and scheduling problem using improved particle swarm optimization. *Soft Computing* **24**(24) 19007–19015.

Pillac, V., C. Gueret, A. L. Medaglia. 2013. A parallel matheuristic for the technician routing and scheduling problem. *Optimization Letters* **7**(7) 1525–1535.

Pillac, V., C. Guéret, A.-L. Medaglia. 2018. A fast reoptimization approach for the dynamic technician routing and scheduling problem. *Recent Developments in Metaheuristics* 347–367.

Pourjavad, E., E. Almehdawe. 2022. Optimization of the technician routing and scheduling problem for a telecommunication industry. *Annals of Operations Research* **315**(1) 371–395.

Savelsbergh, M. W. P. 1992. The vehicle routing problem with time windows: Minimizing route duration. *ORSA Journal on Computing* **4**(2) 146–154.

Sherali, H. D., J. C. Smith. 2001. Improving discrete model representations via symmetry considerations. *Management Science* **47**(10) 1396–1407.

Taillard, É., P. Badeau, M. Gendreau, F. Guertin, J.-Y. Potvin. 1997. A tabu search heuristic for the vehicle routing problem with soft time windows. *Transportation Science* **31**(2) 170–186.

Toth, P., D. Vigo, eds. 2014. *Vehicle Routing: Problems, Methods, and Applications*. 2nd ed. SIAM.

Zamorano, E., R. Stolletz. 2017. Branch-and-price approaches for the multiperiod technician routing and scheduling problem. *European Journal of Operational Research* **257**(1) 55–68.