Journal of Industrial and Management Optimization

JIMO, 22(6): 2878–2911.
DOI: 10.3934/jimo.2026106
Received: 14 January 2026
Revised: 30 April 2026
Accepted: 11 May 2026
Published: 25 May 2026

https://www.aimspress.com/journal/jimo

Research article

Integrated scheduling of drone-based medical sample delivery and testing

Qiuchen Gu[1], Jingyi Chen[2], Bin Hu[1], Qi Zheng[1] and Tijun Fan[3,*]

[1] School of Management, Shanghai University of Engineering Science, Shanghai, China
[2] School of Business, Minjiang University, Fuzhou, China
[3] School of Business, East China University of Science and Technology, Shanghai, China

* **Correspondence:** E-mail: tjfan@ecust.edu.cn.

Abstract: The timeliness of medical sample delivery and testing is critical, particularly within hub-and-spoke networks. Because of their high speed and dispatch flexibility, drones can reduce the total testing completion time (TTCT), defined here as the system-wide makespan, when delivery is coordinated jointly with downstream testing. In this study, the integrated scheduling problem of drone-based medical sample delivery and testing was formulated to minimize the TTCT by synchronizing multi-trip drone scheduling with continuous upstream sample collection and downstream deterministic first-in-first-out (FIFO) queueing at testing institutions. A mixed integer linear programming (MILP) model was developed in standard makespan form to capture the coupling between spatial decisions, including the assignment of the collection point and the testing institution, and temporal decisions, including the dispatch timing, trip frequency, and effective batch size, for an uncertain number of drone trips. A three-stage heuristic framework consisting of the initialized spatial decision, restricted temporal decision, and refined spatial decision was then proposed. The experimental results showed that the heuristic returns solutions within 1.23 s, on average, and reduces the TTCT by 37.60%, on average, relative to a spatial-only comparator.

Keywords: drone delivery; medical sample delivery; integrated scheduling; multi-trip scheduling
Mathematics Subject Classification: 90B06, 90B35, 90C27

1. Introduction

Medical sample testing is a cornerstone of medical diagnoses and treatment plans. Approximately 70% of medical diagnoses are made based on medical test results [1]. To balance cost efficiency with specialized analysis, healthcare systems widely adopt a centralized "hub-and-spoke" model [2], necessitating the frequent delivery of medical samples from dispersed collection points to qualified testing institutions across different geographical locations [3]. Given that delayed test results may lead to compromised medical diagnoses and deferred treatments, the timeliness of medical sample delivery is critical.

Currently, the delivery of medical samples relies predominantly on ground vehicles, yet their efficiency is severely compromised by the road network and traffic unpredictability. Furthermore, ground delivery typically involves low-frequency, large-batch deliveries along fixed routes. This delivery mode creates a mismatch such that continuously collected samples arrive at testing institutions simultaneously in large bulks and trigger workload spikes, leading to queueing congestion and prolonged test waiting times.

Advancements in drone technology offer an effective solution due to the high flight speed of drones and their dispatch flexibility. Zipline's operations in Rwanda demonstrate that drone delivery can reduce not only the in-transit delivery time but also the turnaround time for medical samples by nearly 80% [4]. Unlike ground vehicles, drone delivery enables high-frequency, small-batch deliveries. By dynamically adjusting dispatch timings, batch sizes, and destination testing institutions, sample arrivals can be transformed from a dispersed, large bulk into a continuous sample flow, which smooths workloads and mitigates queueing congestion at downstream testing institutions. Thus, although drones do not participate in testing procedures, their multi-trip scheduling dictates the timeliness of obtaining test results.

Existing drone delivery models focus mainly on optimizing the in-transit delivery time or distance. However, they neglect the downstream testing capacity and queueing congestion. In fact, incorporating the testing stage into the medical sample delivery problem is a structural necessity because the core objective in this context is to minimize the time until testing is completed rather than the in-transit delivery time. Consequently, this study introduces the integrated scheduling problem of drone-based medical sample delivery and testing. The objective is to minimize the total testing completion time (TTCT), which is defined as the system-wide makespan. Within this integration, continuous spatiotemporal scheduling decisions are coupled with a discrete first-in-first-out (FIFO) queueing logic at downstream testing institutions, thus leading to significant complexity in modeling and algorithmic approaches.

The main contributions of this paper are as follows:

(1) We formulate a novel integrated scheduling problem for drone-based medical sample delivery that shifts the optimization paradigm from minimizing the in-transit delivery time to minimizing the system-wide TTCT.

(2) We develop a rigorous MILP model that explicitly characterizes the coupling mechanism between continuous upstream sample collection, multi-trip drone dispatching, and downstream deterministic FIFO queueing dynamics.

(3) To address the computational intractability of the exact MILP model, we propose a highly efficient three-stage heuristic framework that decouples and optimizes complex spatiotemporal variables.

(4) We provide comprehensive computational validations to demonstrate the effectiveness of our proposed modeling and algorithmic approach with spatiotemporal synchronization.

In summary, the research highlights of this study are multifaceted. First, the novel scheduling paradigm shifts the optimization focus from minimizing delivery in-transit time to minimizing TTCT. Second, to capture the delivery–testing mechanism accurately, the proposed MILP formulation explicitly couples multi-trip drone scheduling with the downstream FIFO queueing logic. Third, a three-stage heuristic framework is developed to resolve the NP-hard computational intractability caused by this spatiotemporal coupling in an efficient manner. Lastly, our controlled comparative analyses demonstrate that spatiotemporal synchronization significantly outperforms conventional spatial-only assignment strategies.

The remainder of this paper is organized as follows: Section 2 reviews the related literature. Section 3 formally defines the problem and formulates the optimization model. Section 4 presents the three-stage heuristic framework. Section 5 reports the experimental results of a stylized real-case scenario and 20 generated instances. Section 6 concludes the paper.

2. Literature review

To provide a structured overview of related work, this section is organized into three subsections: the optimization of drone-based delivery systems, the delivery of medical samples, and the integrated scheduling of processing and delivery. This division progressively focuses on drone routing methodologies, the specific timeliness requirements of medical delivery, and the methodological gap with regard to integrating delivery with downstream processing.

2.1. Optimization of drone-based delivery systems

The optimization of drone-based delivery has garnered significant academic attention. The foundational work by Murray and Chu [5] introduced the "flying sidekick traveling salesman problem" (FSTSP), establishing a mathematical framework for coordinating drones with ground vehicles to improve delivery efficiency. Building on this work, Otto et al. [6], Pasha et al. [7], Chung et al. [8], and Dukkanci et al. [9] provided a comprehensive review of drone delivery, highlighting the maturation of optimization methodologies in designing drone-based logistics networks.

Research on drone-based delivery systems focuses on enhancing the routing efficiency of drone delivery or vehicle-drone delivery under realistic operational constraints. Dorling et al. [10] addressed the drone routing problem with battery constraints while considering the cost impacts of battery reuse and recharging. Cheng et al. [11] formulated a drone routing problem with a nonlinear energy function affected by the drone's payload. By integrating weather information, Cheng et al. [12] further extended model realism, utilizing clustering-based fuzzy sets to characterize uncertain flight times and minimize delivery delays. Shi et al. [13] proposed a bi-objective optimization model for the simultaneous pickup and delivery of medical supplies. Zhang et al. [14] optimized synchronized truck–drone delivery in disaster relief by accounting for the priority of different demand sites. Dukkanci et al. [15] proposed a

two-stage stochastic planning approach for post-earthquake relief delivery, emphasizing the robustness of drone operations under uncertainty. Otto et al. [16] investigated delivery policies for truck-and-drone systems in disaster relief. Zhang et al. [17] developed optimization strategies for truck–drone collaborative dispatching in mountainous disaster areas. Wang et al. [18] formulated a truck–drone routing problem with stochastic demand, balancing operating costs with service reliability. Yang et al. [19] studied a vehicle–drone pickup and delivery problem under road congestion, incorporating time-dependent road travel times.

Existing research has established robust methodologies to minimize the total time or cost of drone-based delivery systems. However, these methodologies overlook the operational benefits of high-frequency, small-batch drone deliveries in medical sample delivery with testing. In fact, medical sample delivery must integrate the downstream processing procedure of testing institutions and minimize the total testing completion time (TTCT).

2.2. Delivery of medical samples

The delivery of medical samples is characterized by strict requirements for timeliness and biological integrity. While the fundamental research by Hawkins [20] stated that the turnaround time (TAT) is essential for medical sample tests, Plebani [2] emphasized that the TAT, including the delivery time, is a critical performance indicator for the "hub-and-spoke" vision of medical sample tests. Yücel et al. [21] studied medical sample collection and testing networks to balance testing consolidation and delivery costs. Zabinsky et al. [22] studied the scheduling and routing for medical sample collection from dispersed clinics to a central testing institution to minimize the arrival time of samples at the institution.

More recently, several studies have investigated the drone-based delivery of medical samples. Stierlin et al. [3] reviewed current advancements in this domain, positing that drones are a transformative tool for medical delivery. Kim et al. [23] proposed a drone-aided pickup and delivery network for the delivery of medication and the collection of samples to minimize costs. Chakraborty et al. [24] proposed a drone-assisted two-stage collection system during epidemic outbreaks to expedite the accumulation of samples. Stierlin et al. [25] conducted comparative studies confirming that drone transport preserves the preanalytic integrity of medical samples. Molinari et al. [26] and Zhao et al. [27] further assessed the operational risks and safety intervals of drone flights in complex urban environments. Huang et al. [28] provided empirical evidence from a 5G-enabled medical drone pilot and demonstrated that, compared with ground vehicles, drones could significantly reduce the delivery time for COVID-19 nucleic acid samples.

Although the TAT is recognized as a critical indicator encompassing the entire testing process rather than the delivery phase, optimization studies on sample delivery rarely prioritize it as an objective. Even fewer researchers have integrated the multi-trip scheduling of drones into medical sample delivery to minimize the total testing completion time (TTCT). In fact, for the TTCT, multi-trip drone scheduling significantly influences the in-transit delivery time as well as the queuing and processing time at testing institutions.

2.3. Integrated scheduling of processing and delivery

While the integration of medical sample delivery and testing schedules has received limited attention, the concept of integrated production and distribution scheduling (IPDS) is well established in manufacturing. As reviewed by Chen [29], Moons et al. [30], and Fu et al. [31], IPDS addresses a similar perspective on integrated scheduling with delivery. Hurink and Knust [32] studied job-shop problems with a single transport robot, modeling the transport leg as an operation on a distinct machine. Ullrich [33] demonstrated that jointly optimizing machine operations and vehicle routing is essential for time-sensitive workflows. Hou et al. [34] incorporated explicit delivery time windows into the distributed flow shop scheduling framework. Liang et al. [35] studied the integrated scheduling of production and material delivery that coordinates production lines with automated delivery.

Fewer researchers have studied the integrated scheduling of processing and delivery with drones. Shi et al. [13] proposed a location-routing model for drone delivery, treating multi-trip delivery as being constrained by loading operations at the depot. Zhou et al. [36] studied emergency blood delivery by drones, in which flight delivery is tightly coupled with the drone charging time as a necessary processing constraint. Lejeune and Ma [37] studied the stochastic network design for opioid response, where drone delivery is modeled alongside a service queue with limited processing capacity.

The studies above mainly integrate drone delivery with processing constraints inherent to the drone itself. However, they rarely consider integrating drone delivery with upstream and downstream operational processing. In fact, medical sample delivery necessitates synchronizing continuous sample collection upstream with batch sample testing downstream.

To clarify the methodological gap discussed above and articulate the contributions of this study, Table 1 compares the proposed model with representative related works.

Table 1. Comparison of representative related studies and the proposed model.

Reference	Application context	Multi-trip scheduling	Upstream demand	Downstream processing/queueing	Primary optimization focus
Cheng et al. [11]	General drone routing	No	Static requests	Not modeled	Energy-aware routing
Yücel et al. [21]	Medical specimen collection to a central laboratory	Yes (single vehicle)	Continuous accumulation	Central-facility processing-rate matching	Max processed amount/transport cost
Zabinsky et al. [22]	Medical specimen collection to a central laboratory	No	Predetermined availability times	Not modeled	Pickup-to-lab completion time
Shi et al. [13]	Drone-based medical pickup and delivery	Yes	Static pickup-delivery requests	Depot loading/turnaround	Time-cost-risk routing
Zhou et al. [36]	Emergency blood delivery by drones	No	Static emergency requests	Drone charging constraint	Emergency-delivery scheduling

Continued on next page

Reference	Application context	Multi-trip scheduling	Upstream demand	Downstream processing/queueing	Primary optimization focus
Lejeune and Ma [37]	Opioid-response drone-delivery network	No	Stochastic service requests	Service queue with limited capacity	Average response time
Liang et al. [35]	Integrated production and material delivery	Yes	Production-driven release	Processing-delivery synchronization	Order delays/AGV usage
This paper	Drone-based medical sample delivery and testing	Yes	Continuous accumulation	Deterministic FIFO testing queue	TTCT minimization (system-wide makespan)

As shown in Table 1, the proposed model differs from previous studies in that it explicitly integrates continuous upstream sample accumulation, multi-trip drone dispatching, and downstream deterministic FIFO queueing at testing institutions within a unified optimization framework. The existing related models either focus primarily on routing efficiency or incorporate only partial processing considerations. By contrast, the present study formulates this delivery–testing coupling explicitly in the MILP model, therefore minimizing the total testing completion time (TTCT).

3. Problem definition and model formulation

This section first defines the problem by integrating the delivery–testing setting and the total testing completion time (TTCT) in Section 3.1 and then formulates the optimization model in Section 3.2.

3.1. Problem definition

This study addresses the integrated scheduling problem of drone-based medical sample delivery and testing. In this problem, samples are collected from a series of collection points and delivered to multiple testing institutions by a fleet of drones, and the total testing completion time (TTCT) is minimized. Sample collection constitutes the upstream process, drone transport links the upstream and downstream stages, and laboratory testing is the downstream process. Upon arrival at a testing institution, each delivered batch joins the institution's processing queue.

In the model setting, drones are responsible only for delivering samples from collection points to testing institutions. Drones are not involved in the testing procedure, and tested samples are not assumed to be transported back. Each testing institution is modeled as a deterministic, non-preemptive single-server FIFO processing system.

For the delivery phase executed by drones, multiple drones are initially launched from the drone station. Then, each drone picks up a batch of samples from a collection point, delivers it to a testing institution, and then returns to the drone station to replace its battery in preparation for its next trip. This cycle is defined as a drone trip, as shown in Figure 1. Multi-trip scheduling governs this process

by determining the dispatch timing, batch size, and origin–destination assignment for each drone trip. Such scheduling necessitates an explicit coupling between the spatial and temporal dimensions.

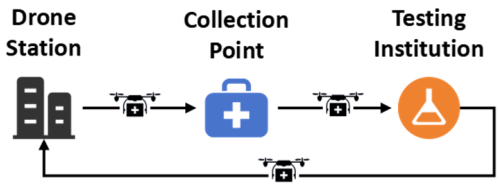

Figure 1. Drone trip for medical sample delivery.

Given the geographically dispersed collection points and testing institutions, the spatial decision optimizes the selection of the origin and destination for each drone trip (e.g., E1 or E2, I1 or I2), as shown in Figure 2. The temporal decision optimizes the dispatch timing for each trip (e.g., l1, l2, and l3), as shown in Figure 3. Collectively, these decisions constitute a schedule that defines the pickup frequency at each collection point. Since samples are collected and accumulate continuously, this frequency inherently dictates the batch size delivered by each drone trip.

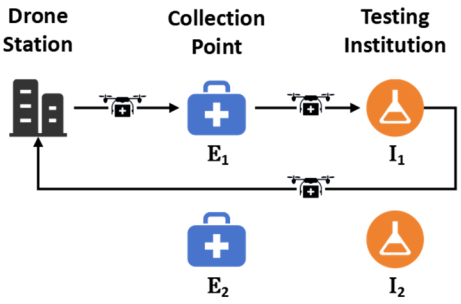

Figure 2. Spatial decision.

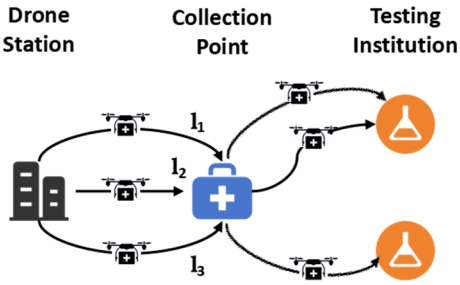

Figure 3. Temporal decision.

In addition, the spatial and temporal decisions jointly determine the receiving times and batch sizes of sequentially arriving sample batches at each testing institution. These arrivals define the workload profile over time and, under the discipline of the deterministic, non-preemptive single-server

FIFO logic, determine the testing completion time of each institution, as illustrated in Figure 4. Therefore, the TTCT is the system's makespan of the integrated delivery–testing process, namely, the maximum testing completion time among all testing institutions after all delivered batches have been processed. The problem seeks to minimize the TTCT by optimizing the multi-trip drone schedule over both the spatial and temporal dimensions.

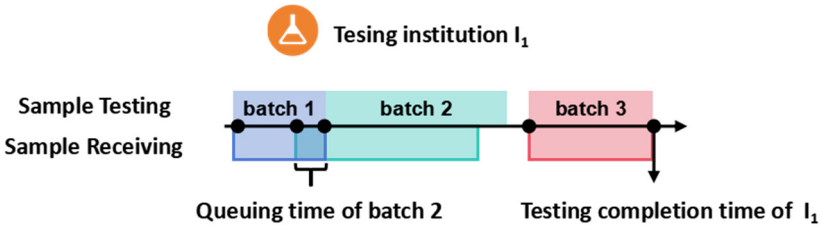

Figure 4. Workload distribution and testing completion time of a testing institution.

3.2. Optimization model

Based on the problem definition above, an optimization model for the integrated scheduling problem of drone-based medical sample delivery and testing is formulated. The optimization model is formulated under the following operating assumptions:

(1) Each trip of a drone picks up samples from one collection point and delivers them to one testing institution.

(2) A fixed handover time is required for the drone to pick up samples from the collection point and transfer them to the testing institution.

(3) A fixed time is required for the drone to replace its battery after each trip.

(4) Each collection point has a fixed sampling speed. However, the sampling speed varies across collection points.

(5) Each testing institution has a fixed testing speed. However, the testing speed varies across testing institutions.

(6) The start and completion times for sample collection at each collection point are known in advance.

(7) All drones are homogeneous. They have the same flight speed, maximum flight time, and maximum load.

To maintain the mathematical tractability of the MILP model, the operational conditions outlined above are intentionally stylized. The constant sampling and testing speeds of assumptions 4 and 5 represent stable expected average values over the operational period. Additionally, physical drone constraints such as energy consumption and battery capacity dynamics are intentionally simplified into conservative deterministic bounds, such as the maximum flight time limit and battery replacement time. The resulting schedule serves as a robust macroscopic baseline for tactical capacity balancing, where minor real-world stochastic fluctuations are naturally buffered by the continuous sample accumulation and downstream queues. Furthermore, external factors, including flight safety and regulatory restrictions, are omitted from this baseline formulation.

The notations of this model are given in Table 2.

Table 2. Notations.

Set	
D	The set of drones, D={1, 2... d... $\|D\|$}
P	The set of trips for each drone, P={0, 1, 2... p... $\|P\|$}
EX	The set of collection points, EX={j_1, j_2...$j_{\|EX\|}$}
INS	The set of testing institutions, INS={k_1, k_2...$k_{\|INS\|}$}
GR	The set of collection points and testing institutions, GR={j_1, j_2...$j_{\|EX\|}$, k_1, k_2...$k_{\|INS\|}$}
V	The set of all nodes, V={0} ∪ GR, where 0 is the drone station
Parameter	
ξ_{p_1p}	Binary, ξ_{p_1p} equals 1 if trip p_1 is executed before trip p by the same drone
c_{ii_1}	The time required for the drone to fly from node i to node i_1
et_j	The sampling start time of collection point j
L_j	The sampling completion time of collection point j
v^1_j	The sampling speed of collection point j, which equals the number of samples that can be collected in a time unit at collection point j
v^2_k	The testing speed of institution k, which equals the number of samples that can be tested in a time unit at testing institution k
$rest$	The time required for the drone to replace its battery after a trip
$take$	The handover time required for the drone to pick up samples at the collection point
$place$	The handover time required for the drone to transfer samples to the testing institution
r	The maximum flight time of the drone
Q	The maximum load of the drone
M	A large positive number
Variable	
B_{dp}	Binary, B_{dp} equals 1 if trip p by drone d exists
t^i_{dp}	The time point at which trip p by drone d departs from node i
s_{dp}	The time point at which trip p by drone d returns to the drone station
E^j_{dp}	Binary, E^j_{dp} equals 1 if trip p by drone d picks up a sample from collection point j
I^k_{dp}	Binary, I^k_{dp} equals 1 if trip p by drone d transfers a sample to testing institution k
q^j_{dp}	The number of samples picked up at collection point j by trip p by drone d
CQ^j_{dp}	The total number of samples picked up at collection point j after the pickup by trip p by drone d
CG^k_{dp}	The time point at which testing institution k can complete its testing after receiving the samples from trip p by drone d
$TTCT$	The system makespan
$\eta^{dp}_{d_1p_1}$	Binary, $\eta^{dp}_{d_1p_1}$ equals 1 if trip p_1 by drone d_1 picks up samples earlier than trip p by drone d at the same collection point
$\mu^{dp}_{d_1p_1}$	Binary, $\mu^{dp}_{d_1p_1}$ equals 1 if trip p_1 by drone d_1 transfers samples earlier than trip p by drone d to the same testing institution

In the notations above, trip $p \in P$ is defined as the p-th trip executed by each drone in chronological order. The length of set P, $|P|$, represents the maximum number of trips that a drone can execute, which is set to the ratio of the longest sample collection time span at the collection points to the shortest drone trip time (rounded up). This approach ensures that the optimization result is not affected.

For simplicity, *(d,p)* pairs are used in the following text to refer to trip p by drone d. Additionally, the model ignores the integrality of the number of samples. Since a single sample typically has a negligible effect on the drone's load, rounding the obtained number of samples to the nearest integer in practical operations does not significantly affect the optimization result.

In accordance with the assumptions and notations presented above, the optimization model of the integrated scheduling problem of drone-based medical sample delivery and testing is formulated. Before the detailed equations are presented, the logical structure of the model is briefly outlined. First, the trip existence and assignment constraints (Eqs. 3–5) govern the activation of drone trips and ensure the unique assignment of collection points and testing institutions. Second, the time consistency constraints (Eqs. 6–14) regulate the chronological sequence of departure, pickup, delivery, base return, maximum flight time limits, and sample availability. Third, the sample capacity and flow constraints (Eqs. 15–21) enforce the drone payload capacities and link the cumulative picked-up quantities to continuous upstream sample generation. Lastly, the linearized queue dynamics constraints (Eqs. 22–36) explicitly translate the downstream deterministic FIFO testing logic into linearized pickup/delivery precedence relationships to evaluate the system-wide completion time. The complete MILP formulation is presented as follows:

$$\min TTCT \tag{1}$$

s.t.

$$TTCT \geq CG_{dp}^{k} \qquad \forall k \in INS, d \in D, p \in P \tag{2}$$

$$B_{dp} \leq B_{dp_1} + M \cdot (1 - \xi_{p_1 p}) \qquad \forall d \in D, p \in P, p_1 \in P \tag{3}$$

$$\sum_{j \in EX} E_{dp}^{j} = B_{dp} \qquad \forall d \in D, p \in P \tag{4}$$

$$\sum_{k \in INS} I_{dp}^{k} = B_{dp} \qquad \forall d \in D, p \in P \tag{5}$$

$$t_{dp}^{0} \geq s_{dp_1} + rest - M \cdot (2 - B_{dp} - \xi_{p_1 p}) \qquad \forall d \in D, p \in P, p_1 \in P \tag{6}$$

$$t_{dp}^{j} \geq t_{dp}^{0} + take + c_{0j} - M \cdot (1 - E_{dp}^{j}) \qquad \forall j \in EX, d \in D, p \in P \tag{7}$$

$$t_{dp}^{k} \geq t_{dp}^{j} + place + c_{0j} - M \cdot (2 - E_{dp}^{j} - I_{dp}^{k}) \qquad \forall j \in EX, k \in INS, d \in D, p \in P \tag{8}$$

$$s_{dp} \geq t_{dp}^{k} + c_{k0} - M \cdot (1 - I_{dp}^{k}) \qquad \forall k \in INS, d \in D, p \in P \tag{9}$$

$$s_{dp} \leq t_{dp}^{k} + c_{k0} + M \cdot (1 - I_{dp}^{k}) \qquad \forall k \in INS, d \in D, p \in P \tag{10}$$

$$s_{dp} - t^0_{dp} \le r \qquad \forall d \in D,\ p \in P \tag{11}$$

$$t^0_{d1} \ge 0 \qquad \forall d \in D \tag{12}$$

$$t^i_{dp} \le M \cdot B_{dp} \qquad \forall i \in GR,\ d \in D,\ p \in P \tag{13}$$

$$t^j_{dp} \ge et_j - M \cdot (1 - E^j_{dp}) \qquad \forall j \in EX,\ d \in D,\ p \in P \tag{14}$$

$$q^j_{dp} \ge 0 \qquad \forall j \in EX,\ d \in D,\ p \in P \tag{15}$$

$$q^j_{dp} \le Q \qquad \forall j \in EX,\ d \in D,\ p \in P \tag{16}$$

$$\sum_{\substack{d \in D \\ p \in P}} q^j_{dp} = (L_j - et_j) \cdot vs_j \qquad \forall j \in EX \tag{17}$$

$$CQ^j_{dp} \ge q^j_{dp} - M \cdot \sum_{\substack{d_1 \in D \\ p_1 \in P}} \eta^{dp}_{d_1 p_1} \qquad \forall j \in EX,\ d \in D,\ p \in P \tag{18}$$

$$CQ^j_{dp} \ge CQ^j_{d_1 p_1} + q^j_{dp} - M \cdot (1 - \eta^{dp}_{d_1 p_1}) \qquad \forall j \in EX,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{19}$$

$$CQ^j_{dp} \le (t^j_{dp} - st_j) \cdot v^1{}_j \qquad \forall j \in EX,\ d \in D,\ p \in P \tag{20}$$

$$q^j_{dp} \le M \cdot E^j_{dp} \qquad \forall j \in EX,\ d \in D,\ p \in P \tag{21}$$

$$\eta^{dp}_{d_1 p_1} \ge (t^j_{dp} - t^j_{d_1 p_1}) / M - M \cdot (1 - E^j_{d_1 p_1}) \qquad \forall j \in EX,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{22}$$

$$\eta^{dp}_{d_1 p_1} \le 1 - (t^j_{d_1 p_1} - t^j_{dp}) / M \qquad \forall j \in EX,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{23}$$

$$\eta^{dp}_{d_1 p_1} \le E^j_{d_1 p_1} + M \cdot \left(1 - E^j_{dp}\right) \qquad \forall j \in EX,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{24}$$

$$\eta^{dp}_{d_1 p_1} \le E^j_{dp} + M \cdot \left(1 - E^j_{d_1 p_1}\right) \qquad \forall j \in EX,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{25}$$

$$\eta^{d_1 p_1}_{dp} + \eta^{dp}_{d_1 p_1} \ge 1 - M \cdot (2 - E^j_{dp} - E^j_{d_1 p_1}) \qquad \forall j \in EX,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{26}$$

$$\eta^{d_1 p_1}_{dp} + \eta^{dp}_{d_1 p_1} \le 1 + M \cdot (2 - E^j_{dp} - E^j_{d_1 p_1}) \qquad \forall j \in EX,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{27}$$

$$CG^k_{dp} \ge t^k_{dp} + q^j_{dp} / vt_k - M \cdot \sum_{\substack{d_1 \in D \\ p_1 \in P}} \mu^{dp}_{d_1 p_1} \qquad \forall k \in INS,\ d \in D,\ p \in P \tag{28}$$

$$CG^k_{dp} \ge CG^k_{d_1 p_1} + q^j_{dp} / v^2{}_k - M \cdot (1 - \mu^{dp}_{d_1 p_1}) \qquad \forall j \in EX,\ k \in INS,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{29}$$

$$CG^k_{dp} \ge t^k_{dp} + q^j_{dp} / v^2{}_k \qquad \forall j \in EX,\ k \in INS,\ d \in D,\ p \in P \tag{30}$$

$$\mu^{dp}_{d_1 p_1} \ge (t^k_{dp} - t^k_{d_1 p_1}) / M - M \cdot (1 - I^k_{d_1 p_1}) \qquad \forall k \in INS,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{31}$$

$$\mu^{dp}_{d_1 p_1} \le 1 - (t^k_{d_1 p_1} - t^k_{dp}) / M \qquad \forall k \in INS,\ d \in D,\ p \in P,\ d_1 \in D,\ p_1 \in P \tag{32}$$

$$\mu_{d_1p_1}^{dp} \le I_{d_1p_1}^{k} + M \cdot \left(1 - I_{dp}^{k}\right) \qquad \forall k \in INS, d \in D, p \in P, d_1 \in D, p_1 \in P \tag{33}$$

$$\mu_{d_1p_1}^{dp} \le I_{dp}^{k} + M \cdot \left(1 - I_{d_1p_1}^{k}\right) \qquad \forall k \in INS, d \in D, p \in P, d_1 \in D, p_1 \in P \tag{34}$$

$$\mu_{d_1p_1}^{dp} + \mu_{dp}^{d_1p_1} \ge 1 - M \cdot (2 - I_{dp}^{k} - I_{d_1p_1}^{k}) \qquad \forall k \in INS, d \in D, p \in P, d_1 \in D, p_1 \in P \tag{35}$$

$$\mu_{d_1p_1}^{dp} + \mu_{dp}^{d_1p_1} \le 1 + M \cdot (2 - I_{dp}^{k} - I_{d_1p_1}^{k}) \qquad \forall k \in INS, d \in D, p \in P, d_1 \in D, p_1 \in P \tag{36}$$

The objective function (1) minimizes the system's makespan TTCT. Combined with constraint (2), it denotes the latest testing completion time among all testing institutions.

Trip existence and assignment. Constraints (3)–(5) represent the trip existence and assignment constraints. Constraint (3) ensures the coherence of drone trips such that the existence of trip *(d,p)* is contingent upon the existence of trip *(d,p₁)* when $p_1<p$. Constraints (4) and (5) ensure that each existing trip is assigned to exactly one collection point and exactly one testing institution.

Time consistency. Constraints (6)–(14) represent the time consistency constraints. Constraint (6) ensures that a drone can start a trip only after it returns from any previous trip and completes battery replacement. Constraint (7) links departure from the drone station to departure from the collection point by adding the flight time and pickup handover time. Constraint (8) links departure from the collection point to departure from the testing institution by adding the flight time and delivery handover time. Constraints (9) and (10) define the return time of each trip to the drone station. Constraint (11) enforces the maximum flight time. Constraint (12) ensures a nonnegative departure time for the first trip by each drone. Constraint (13) activates the key time points only when a trip exists. Constraint (14) ensures that pickup occurs after sample collection starts at the corresponding collection point.

Sample capacity and flow. Constraints (15)–(21) represent the sample capacity and flow constraints. Constraint (15) enforces nonnegative pickup quantities. Constraint (16) limits the pickup quantity of each trip according to the drone's capacity. Constraint (17) ensures that all samples collected at each collection point are eventually picked up. Constraints (18) and (19) define the cumulative number of picked-up samples at the same collection point according to the pickup order. Constraint (20) ensures that the cumulative picked-up samples do not exceed the number collected by the pickup time. Constraint (21) links the pickup quantity to the assignment of the collection point within a trip.

Linearized queue dynamics. Constraints (22)–(36) represent the linearized queue dynamics constraints. Constraints (22)–(27) are the linearized form of Eq. (37).

$$\eta_{d_1p_1}^{dp} = \begin{cases} 1 & if \quad \sum_{j \in EX} E_{dp}^{k} \cdot E_{d_1p_1}^{k} = 1 \quad and \quad t_{dp}^{j} > t_{d_1p_1}^{j} \\ 0 & else \end{cases} \qquad \forall d \in D, p \in P, d_1 \in D, p_1 \in P \tag{37}$$

Constraint (22), combined with constraint (13), states that $\eta_{d_1p_1}^{dp}=1$ if trip *(d, p)* and trip *(d₁, p₁)* pick up samples at the same collection point, while trip *(d, p)* picks up samples later than trip (d_1, p_1) at that collection point. Constraint (23), combined with constraint (13), states that $\eta_{d_1p_1}^{dp}=0$ if trip *(d, p)* and trip *(d₁, p₁)* pick up samples at the same collection point, while trip *(d, p)* picks up samples earlier than trip *(d₁, p₁)* at that collection point. Constraints (24) and (25) state that $\eta_{d_1p_1}^{dp}=0$ if trip *(d, p)* and trip *(d₁, p₁)* do not pick up samples from the same collection point. Constraints (26) and (27) provide supplementary clarification for the special case in which trip *(d, p)* and trip *(d₁, p₁)* pick up samples at exactly the same time and the same collection point.

Constraints (28)–(30) ensure that the testing completion time at each testing institution is no earlier than both the completion time of previously received samples plus the testing time of the current batch and the delivery time of the current batch plus its testing time. Constraints (31)–(36) are the linearized form of Eq. (38). The linearization is performed in the same manner as that of constraints (22)–(27).

$$\mu_{d_1p_1}^{dp} = \begin{cases} 1 & if \quad \sum_{k \in INS} I_{dp}^{k} \cdot I_{d_1p_1}^{k} = 1 \ \ and \ \ t_{dp}^{k} > t_{d_1p_1}^{k} \\ 0 & else \end{cases} \qquad \forall d \in D, p \in P, d_1 \in D, p_1 \in P \tag{38}$$

Figure 5 illustrates a two-trip, one-institution example of η and μ. At collection point E1, trip (1,1) reaches the point earlier than trip (2,1); thus, $\eta_{1,1}^{2,1} = 1$. At testing institution I1, trip (1,1) also arrives earlier than trip (2,1); thus, $\mu_{1,1}^{2,1} = 1$. Accordingly, the quantity carried by trip (1,1) is counted before that of trip (2,1) in the cumulative pickup tracking, and the testing completion time associated with trip (2,1) must be propagated after that of trip (1,1) under the FIFO service.

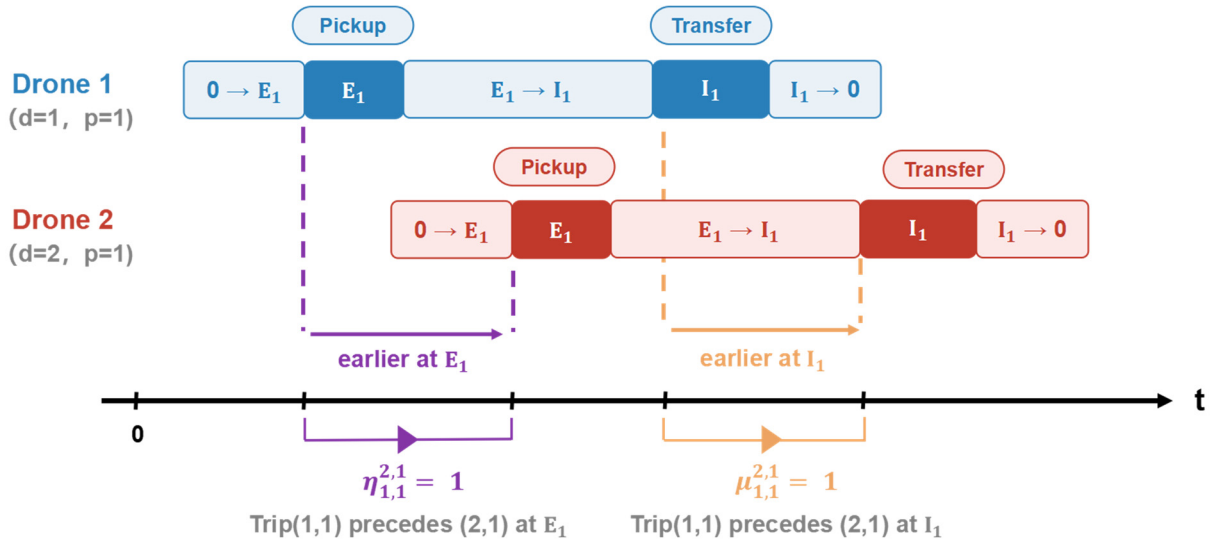

Figure 5. Timeline illustration of the precedence variables η and μ.

4. Algorithms

The integrated scheduling problem of drone-based medical sample delivery and testing by drones requires spatial and temporal decisions for uncertain numbers of drone trips. Owing to its complexity, a direct exact solution becomes computationally challenging for practical-scale instances. Therefore, a heuristic framework that includes the initialized spatial decision, restricted temporal decision, and refined spatial decision is introduced in this section.

Within this framework, the restricted temporal decision is pivotal. As shown in Figure 6, the timing of a single trip cannot be optimized independently because it depends on the dispatch timing of preceding trips. We therefore jointly optimize the dispatch timing and delivery frequency of trips serving the same collection point under the temporary restriction that all such trips are executed by the same drone and are delivered to the same testing institution.

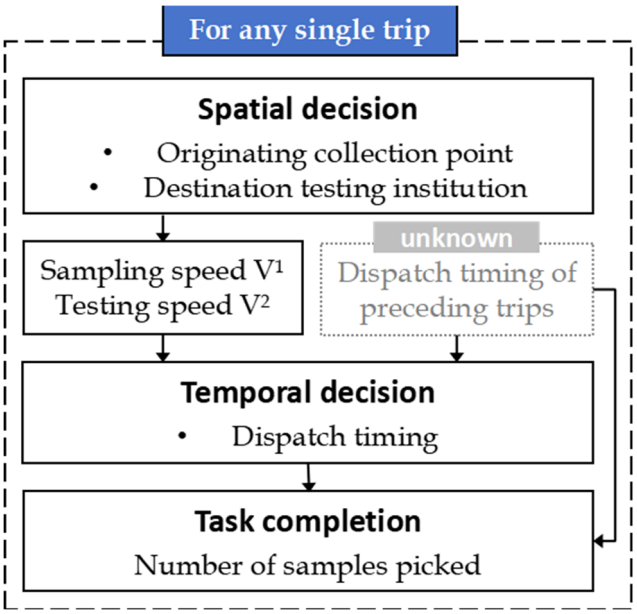

Figure 6. Determination strategy of a single trip.

Under this restriction, the initialized spatial decision allocates one testing institution to each collection point. The restricted temporal decision then schedules the trips executed by one drone to deliver all samples from that collection point to the assigned institution. Finally, the refined spatial decision relaxes this restriction by adjusting the trip destinations to further improve the global solution. The overall framework is shown in Figure 7.

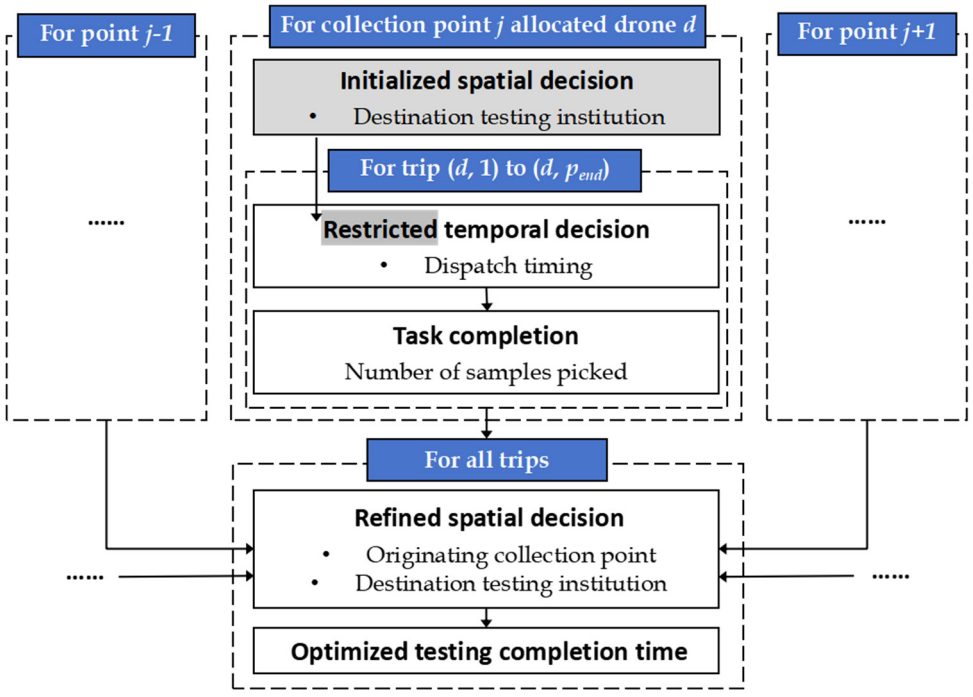

Figure 7. Structure of the framework.

The novelty of the proposed framework lies in its spatial–temporal decoupling and synchronization mechanism, which translates downstream dynamic queueing states at testing institutions into coordinated temporal scheduling and destination-adjustment decisions across the three stages.

To facilitate the presentation from the outset, Table 3 summarizes the notation used in the restricted temporal decision. Because it is solved for a collection point already associated with one drone and one testing institution, the subscripts for drones, collection points, and testing institutions in Table 1 are omitted here for simplicity.

Table 3. Notations for the restricted temporal decision.

Parameter		
c		The time required for the drone to fly from the collection point to the testing institution, plus the handover time at both places
cir_time		The total time consumed for a whole trip
L		The sampling completion time of the collection point
v^1		The sampling speed of the collection point
v^2		The testing speed of the testing institution
Q		The maximum load of drone
G		$G = v^1 * L$
		Total number of samples collected at the collection point within its active time
Decision variable		
q_p	$\forall p \in P$	The number of samples picked up at the collection point by trip p
l_p	$\forall p \in P$	The time point at which trip p starts to pick up samples at the collection point
Dependent variable		
at_p	$\forall p \in P$	$at_p = l_p + c$
st_p	$\forall p \in P$	The time point at which trip p finishes transferring samples to the institution $st_p = \max\{at_p, ft_{p-1}\}$
ft_p	$\forall p \in P$	The time point at which the testing of samples transferred by trip p can start $ft_p = st_p + q_p / v^2$
		The time point at which the testing of samples transferred by trip p completes

With the notations in Table 3 established in advance, Algorithm 1 presents the complete pseudo-code of the three-stage heuristic framework.

Algorithm 1: Three-stage heuristic framework

Input: Sets EX, INS, D; locations loc_dep, loc_ex, loc_ins; speeds v^1, v^2; capacity Q; sampling completion times L

Output: Optimized schedule $S^* = \{\text{trip assignments, dispatch timings, batch sizes}\}$ that minimizes TTCT

1 // Precompute distance parameters

2 for each ($j \in EX, k \in INS$) do

3 $cir\_time(j,k) \leftarrow c_{0,j} + c_{j,k} + c_{k,0}$

4 $dis\_ex\_ins(j,k) \leftarrow c_{j,k}$

5 end for

6
7 // === Stage 1: Initialized Spatial Decision ===
8 $X^* \leftarrow \text{SolveAllocationModel}(EX, INS, v^1, v^2, L, dis\_ex\_ins)$
9 // $X^*(j)$ returns the testing institution allocated to collection point$j$
10
11 // === Stage 2: Restricted Temporal Decision ===
12 for each $k \in INS$ do
13 $EX_k \leftarrow \{j \in EX: X^*(j) = k\}$ // collection points allocated to k
14 if $|EX_k| = 0$ then continue
15 for each $j \in EX_k$ do
16 $\{l_p, q_p\}_{p=1..n} \leftarrow \text{RestrictedTemporalDecision}(j, k, v_j^1, v_k^2, cir\_time(j,k), Q, L_j, |EX_k|)$
17 Store trip records: (j, k, l_p, q_p) for each trip$p$
18 end for
19 $ft_k \leftarrow \text{CalcTestingCompletionTime}(\text{all}k, v_k^2)$
20 end for
21
22 // === Stage 3: Refined Spatial Decision ===
23 $S^* \leftarrow \text{RefinedSpatialDecision}(\text{trip}, ft, v^2, cir\_time, dis\_ex\_ins)$
24
25 returnS^*

Based on the overall framework and notation introduced above, Section 4.1 develops the restricted temporal decision under the temporary restriction, Section 4.2 presents the initialized and refined spatial decision procedures, and Section 4.3 reports the computational complexity of the three-stage framework.

4.1. The restricted temporal decision approach

Using the notations introduced at the beginning of Section 4, this subsection first derives the optimality conditions and then presents the heuristic routine for the restricted temporal decision.

(1) Optimality conditions of the restricted temporal decision

Assuming that the current collection point *j* is allocated an idle drone *d*, the delivery frequency and dispatch timing for trips (*d, p1*) to (*d, pend*) are determined subject to the following key constraints:

1) $l_n \geq L$, where $n=p_{end}$, to ensure that no trip picks up samples collected within (l_n, L) at j.

2) $l_{p+1} - l_p \geq cir\_time$, to constrain the minimum time interval between the dispatch timing of each pair of consecutive trips.

3) $q_p \leq Q$, to constrain the maximum number of samples that are picked up by a single trip.

In addition to the constraints above, there is an implicit constraint that can ensure the local optimality of the restricted temporal decision:

4) $q_p = (l_{p+1} - l_p) \cdot v^1$ if $cir\_time \cdot v^1 \leq Q$. Combined with $q_p \leq Q$, it also holds true that $l_{p+1} - l_p \leq Q/v^1$ if $cir\_time \cdot v^1 \leq Q$. This constraint limits both the time interval between consecutive trips and the number of samples picked up by each trip to eliminate the solutions with deliberately delayed or underloaded trips.

With the constraints above, the objective of the restricted temporal decision is

$$\min\ ft_n \tag{40}$$

which minimizes the testing completion time of samples transferred from the last trip of drone *d*, in which ft_n can be equivalently expressed as equations (41) and (42).

$$ft_n = at_n + \max\{0, ft_{n-1} - at_n\} + q_n / v^2 \tag{41}$$

$$ft_n = at_1 + G/v^2 + \max\left\{0, \sum_{i \in P}(at_i - ft_{i-1})\right\} \tag{42}$$

Equation (41) equals the arrival time of the last batch plus the remaining testing time of previously arrived samples and the testing time of the last batch. The term is the backlog time. Although it is generally difficult to compute, equation (41) becomes easy to evaluate when the backlog time is guaranteed to be zero.

Equation (42) equals the arrival time of the first batch plus the testing time of all samples collected at j and the idle time accumulated during the testing horizon. The term is the total idle time between the completion of all available tests and the arrival of the next batch. When this idle time is guaranteed to be zero, equation (42) is easy to evaluate.

Accordingly, the following lemmas characterize the optimality conditions for the cases:

Lemma 1: If $v^1 \le v^2$ and $cir\_time \cdot v^1 \le Q$, the optimal solution of the restricted temporal decision satisfies the following:

$$\hat{l}_n = L \tag{43}$$

$$\widehat{l_{n-1}} = L - cir\_time \tag{44}$$

The minimized objective function value is as follows:

$$\widehat{ft_n} = L + c + cir\_time \cdot v^1 / v^2 \tag{45}$$

Lemma 2: If $v^1 > v^2$ and $cir\_time \cdot v^1 \le Q$, the optimal solution of the restricted temporal decision satisfies the following:

$$\hat{l}_1 = cir\_time \cdot v^2 / v^1 \tag{46}$$

$$\hat{l}_2 = cir\_time \cdot (1 + v^2 / v^1) \tag{47}$$

The minimized objective function value is as follows:

$$\widehat{ft_n} = cir\_time \cdot v^2 / v^1 + c + G/v^2 \tag{48}$$

The proofs of the lemmas above are addressed in detail in the Appendix.

(2) Heuristic approach for the restricted temporal decision

Algorithm 2 provides the restricted temporal decision routine used in Stage 2 of the heuristic framework.

Algorithm 2: Restricted temporal decision

Input: Collection point j, institution k, speeds v_j^1, v_k^2, $cir\_time(j,k)$, capacity Q, sampling completion time L_j, number of collection points allocated to k: N_k

Output: Dispatch timings $\{l_p\}$ and batch sizes $\{q_p\}$ for trips $p = 1, \ldots, n$

$G \leftarrow L_j \cdot v_j^1$ // total samples to collect
$c \leftarrow dis\_ex\_ins(j,k)$
if $cir\_time \cdot v_j^1 \le Q$ then // Case A
    if $v_j^1 \le v_k^2$ then // Lemma 1
        $l_n \leftarrow L_j; q_n \leftarrow cir\_time \cdot v_j^1$
        Generate trips backward: $l_p \leftarrow l_{p+1} - cir\_time, q_p \leftarrow cir\_time \cdot v_j^1$
    else // Lemma 2: $v_j^1 > v_k^2$
$l_1 \leftarrow cir\_time \cdot v_k^2 / v_j^1; q_1 \leftarrow cir\_time \cdot v_k^2$
        while $\sum q_p < G$ do
$l_p \leftarrow l_{p-1} + cir\_time$
$q_p \leftarrow \min(cir\_time \cdot v_j^1, G - \sum q)$
else // Case B: $cir\_time \cdot v_j^1 > Q$
    $l_1 \leftarrow Q / v_j^1; q_1 \leftarrow Q$
    while $\sum q_p < G$ do
        $l_p \leftarrow l_{p-1} + cir\_time; q_p \leftarrow \min(Q, G - \sum q)$
// Adjustment for multiallocation
if $N_k > 1$ and $v_j^1 > v_k^2$ and $cir\_time \cdot v_j^1 \le Q$ then
    $q_1 \leftarrow Q \cdot v_k^2 / v_j^1; l_1 \leftarrow Q \cdot v_k^2 / (v_j^1)^2$
    Regenerate subsequent trips with $l_p \leftarrow l_{p-1} + cir\_time, q_p \leftarrow \min(Q, G - \sum q)$
return $\{l_p, q_p\}_{p=1}^n$

Assuming that the current collection point j is allocated an idle drone d to deliver samples to testing institution k, if $cir\_time \cdot v^1 \le Q$ is satisfied, the restricted temporal decision will be made according to the optimality conditions addressed in part (1).

If $cir\_time \cdot v^1 \le Q$ and $v^1 \le v^2$, the optimal objective function value is as shown in equation (45), which can be obtained by multiple strategies as long as $ft_{n-1} \le at_n$ is satisfied. The following strategy is adopted in the heuristic approach of this study:

$$l_n = L \tag{49}$$

$$q_n = cir\_time \cdot v^1 \tag{50}$$

$$l_{p-1} = \max\{0, l_p - cir\_time\} \qquad \forall p \in P \tag{51}$$

$$q_{p-1} = \max\{0, q_p - cir\_time \cdot v^1\} \qquad \forall p \in P, p > 1 \tag{52}$$

Provided that $v^1 \le v^2$, testing institution *k* is more likely to have idle time than backlog time; thus, the subsequent refined spatial decision tends to redirect other trips to *k*. In the strategy, as many trips as possible are generated to create a sufficient "gap" for such redirection.

If $cir\_time \cdot v^1 \le Q$ and $v^1 > v^2$, the optimal objective function value is as shown in equation (48), which can also be obtained by multiple strategies as long as $l_{p+1} - l_p \le \left(l_p - l_{p-1}\right) \cdot v^1 / v^2$ is satisfied. The following strategy is adopted in the heuristic approach of this study:

In case $v^1 > v^2$,

$$l_1 = cir\_time \cdot v^2 / v^1 \tag{53}$$

$$q_1 = cir\_time \cdot v^2 \tag{54}$$

$$l_{p+1} = \min\{L, l_p + \min\{Q, (l_p - l_{p-1}) \cdot v^1 / v^2\}\} \qquad \forall p \in P, p < n \tag{55}$$

$$q_p = (l_p - l_{p-1}) \cdot v^1 \quad \forall p \in P, p > 1 \tag{56}$$

institution k is more likely to exhibit backlog time. Therefore, the heuristic differentiates trip intervals and batch sizes so that the backlog can be more finely balanced across institutions in the subsequent refined spatial decision.

However, the optimality conditions addressed in part (1) are unavailable if $cir\_time \cdot v^1 > Q$. In this case, the heuristic simply sets $l_{p+1} - l_p = cir\_time$ and $q_p = Q$ to avoid deliberate delay and underload as follows:

$$q_1 = Q \tag{57}$$

$$l_1 = Q / v^1 \tag{58}$$

$$q_p = \min\{Q, G - \sum_{i \in P, i < p} q_i\} \quad \forall p \in P \tag{59}$$

$$l_{p+1} = l_p + cir\_time \quad \forall p \in P, p < n \tag{60}$$

4.2. The spatial decision approach

The initialized spatial decision assigns each collection point to a testing institution before detailed temporal scheduling is performed. The refined spatial decision then improves this initial allocation by redirecting selected trips to rebalance the downstream workload. In practice, the redirection step is most valuable when the testing workloads are strongly imbalanced across institutions.

(1) Heuristic approach for initialized spatial decision

The initialized spatial decision is to link each collection point to one testing institution, which obtains parameters such as *cir_time* and v^2 for the restricted temporal decision. This decision essentially involves the problem of allocating testing institutions to collection points. Unlike traditional allocation problems, since the objective of this study is to minimize the testing completion time instead of the in-transit delivery time, balancing the testing workload across different testing institutions and

reducing the in-transit delivery time are equally important. Therefore, an allocation problem model with workload balance is formulated for the initialized spatial decision:

$$\min \lambda \cdot \sum_{j\in EX,\, k\in INS} X_{jk} + (1-\lambda)\cdot div\_F \tag{61}$$

s.t.

$$\sum_{k\in INS} X_{jk} = 1 \qquad \forall j \in EX \tag{62}$$

$$F_k = \sum_{j\in EX} X_{jk}\cdot L_j \cdot V_j^1 \big/ V_k^2 \qquad \forall k \in INS \tag{63}$$

$$div\_F = DF(F_k), k \in INS \tag{64}$$

Objective (61) is the linearization of the bi-objective that minimizes the in-transit delivery time and the difference in testing workload time among testing institutions. Constraint (62) ensures that each collection point is allocated one testing institution. Constraint (63) is used to calculate the testing workload time F_k for each testing institution. Constraint (64) is used to measure the difference in F_k across k by the function $DF(F_k)$, which can be rational for adopting a variance or range of F_k.

The model can be solved exactly or heuristically, depending on the problem size. Once the allocation is obtained, each testing institution maintains a list of assigned collection points, and the restricted temporal decision is then solved according to the size of this list.

a) If a testing institution is assigned to only one collection point, the restricted temporal decision in Section 4.1 directly determines the trip timings and batch sizes, from which its testing completion time is obtained.

b) If the testing institution is not allocated to any collection point, no operation will be executed at this initialized spatial decision stage.

c) If the testing institution is allocated to multiple collection points, for those collection points satisfying $cir\_time\cdot v^1 > Q$ and $v^1 > v^2$, the restricted temporal decision is adjusted as follows:

$$q_1 = Q\cdot v^2 \big/ v^1 \tag{65}$$

$$l_1 = Q\cdot v^2 / (v^1)^2 \tag{66}$$

$$q_p = \min\{Q, G - \sum_{i\in P, i<p} q_i\} \quad \forall p \in P \tag{67}$$

$$l_{p+1} = l_p + cir\_time \quad \forall p \in P, p < n \tag{68}$$

This adjustment advances the dispatch timing of the first trip from the collection point. The reason is that when a testing institution is allocated to only one collection point, condition $cir\_time\cdot v^1 > Q$ may result in idle time for this testing institution even though $v^1 > v^2$. Such idle time is more likely to be eliminated when a testing institution is allocated to multiple collection points. Thus, advancing the dispatch timing of the first trip can help the testing institution complete the testing earlier rather than merely increasing idle time.

For the remaining collection points in situation 3), the restricted temporal decision is applied as in Section 4.1. The testing completion time of the institution is then obtained by integrating the dispatch schedules of all trips originating from its assigned collection points.

(2) Heuristic approach for the refined spatial decision

The refined spatial decision relaxes the initial allocation so that samples can be delivered to different institutions at different time points. This decision keeps the dispatch timing and pickup quantity of each trip fixed from the restricted temporal decision and improves the schedule only by redirecting destinations, thereby transferring the workload from institutions with backlog time to those with idle time.

Algorithms 3 and 4 present the refined spatial decision heuristic and the FIFO testing completion routine used to update the workloads of institutions under the FIFO service.

Algorithm 3: Refined spatial decision (local improvement heuristic)

Input: Trip records from Stage 2; testing completion times $\{\mathrm{ft_k}\}$; speeds v^2; distances cir_time,dis_ex_ins

Output: Refined schedule $\mathcal{S}^*$ with minimized TTCT

$\mathcal{S}_{\text{best}} \leftarrow$ current schedule; $MG_{\text{best}} \leftarrow \max_k\{ft_k\}$

for $iter = 1$ to N_{restart} do

  Reset to initial schedule from Stage 2

  $improved \leftarrow$ true

  for $t = 1$ to $T_{\max}$ do

    // Step 1: Select source and target institutions

    if $t \le T_{\text{det}}$ and $improved$ then

      $k_{\text{src}} \leftarrow arg\max_k\{ft_k\}$; $k_{\text{tar}} \leftarrow arg\min_k\{ft_k\}$

    else // probabilistic phase

      $k_{\text{src}} \leftarrow$ RouletteSelect($\{ft_k\}$, exclude min)

      $k_{\text{tar}} \leftarrow$ RouletteSelect($\{ft_k : ft_k < ft_{k_{\text{src}}}\}$, inverse weighting)

    $improved \leftarrow$ false; $best\_gain \leftarrow -\infty$

    // Step 2: Evaluate each trip for redirection

    for each trip s assigned to k_{src} do

      $\Delta_{\text{src}} \leftarrow ft_{k_{\text{src}}} -$CalcTestingCompletionTime (trips of $k_{\text{src}} \setminus \{s\}$)

      $delay_{\text{src}} \leftarrow$ cascade delay on successor trips of same drone

      Recalculate arrival time of s at k_{tar}

      $\Delta_{\text{tar}} \leftarrow$ CalcTestingCompletionTime (trips of $k_{\text{tar}} \cup \{s\}$) $- ft_{k_{\text{tar}}}$

      $gain \leftarrow \Delta_{\text{src}} - \Delta_{\text{tar}} - delay_{\text{src}}$

      if $gain > best\_gain$ and $\max(\text{updated}ft) \le \max(\text{current}ft)$ then

        $best\_gain \leftarrow gain$; record redirect$(s, k_{\text{src}} \to k_{\text{tar}})$

        $improved \leftarrow$ true

    // Step 3: Apply best redirect

    if $improved$ then

      Execute redirect; update $ft_{k_{\text{src}}}$ and $ft_{k_{\text{tar}}}$

31 if $\max_k\{ft_k\} < MG_{\text{best}}$ then
32 $MG_{\text{best}} \leftarrow \max_k\{ft_k\}$; $\mathcal{S}_{\text{best}} \leftarrow$ current schedule
33
34 return $\mathcal{S}_{\text{best}}$

To accurately evaluate the downstream workload, each testing institution is explicitly modeled as a deterministic, non-preemptive first-in-first-out (FIFO) single-server queue. Once a sample batch begins testing, its processing cannot be interrupted, and subsequent batches must wait in the queue, as formalized in Algorithm 4.

Algorithm 4: Calculation of the testing completion time

Input: Sorted arrival-batch pairs $\{(\mathrm{at_i}, \mathrm{q_i})\}_{\mathrm{i=1}}^{\mathrm{m}}$ at institution k (sorted by $\mathrm{at_i}$); testing speed$\mathrm{v_k^2}$
Output: Testing completion time ft
1 $st_1 \leftarrow at_1$ // testing starts when first batch arrives
2 $ft_1 \leftarrow st_1 + q_1/v_k^2$
3 for $i = 2$ to m do
4 $st_i \leftarrow \max(ft_{i-1}, at_i)$ // FIFO queuing
5 $ft_i \leftarrow st_i + q_i/v_k^2$
6 return ft_m

When a trip is redirected to an institution that is farther away, the drone returns to the station later, which can delay all successor trips of the same drone and, in turn, the completion time of the institution originally served by those trips. This cascading effect is explicitly evaluated when deciding whether a candidate redirection should be accepted.

4.3. Computational complexity for the three-stage heuristic framework

Let $n = |EX|$, $m = |INS|$, and N denote the total number of drone trips generated across all collection points. In Stage 1, the allocation model contains $O(nm)$ binary variables and $O(n + m)$ constraints, which can be solved efficiently by commercial solvers for practical-sized instances. Stage 2 generates trip schedules for all collection points in $O(N)$ time and computes the FIFO testing completion times by sorting and scanning the assigned trips at each institution, yielding an overall complexity of $O(N\,log\,N)$. Stage 3 dominates the total runtime. In each iteration, the algorithm evaluates at most $O(N/m)$ trips from the source institution and recomputes the FIFO completion times for the source and target institutions in $O(N/m)$ time per evaluation. With R restarts and at most I iterations per restart, the worst-case complexity of Stage 3 is $O(R \cdot I \cdot (N/m)\text{^}2)$. In our experiments, $R = 20$, $I = 100$, and the roulette-wheel selection in the probabilistic phase of Stage 3 uses a fixed random seed of 42. The early termination rule in the probabilistic phase typically reduces the actual number of iterations well below I. Therefore, the heuristic exhibits polynomial-time behavior and is expected to scale more gracefully than the exact MILP.

5. Experimental study and computational analysis

This section reports two complementary computational studies. Section 5.1 uses a stylized real-case scenario to validate the MILP model of Section 3 through exact-solver analysis and sensitivity tests of key parameters. Section 5.2 then evaluates the heuristic framework of Section 4 on larger randomly generated instances, reports the solution quality and computational time, and compares the proposed approach with a spatial-only nearest-institution comparator under the same operating setting.

All algorithms were implemented in Python, and the MILP benchmark runs were solved with Gurobi. The experiments were conducted on a desktop computer equipped with an Intel(R) Core(TM) i9-14900K 3.20 GHz processor and 32 GB of RAM under a 64-bit operating system.

5.1. Stylized real-case scenario: Exact-solver validation and sensitivity analysis

A stylized real-case scenario motivated by emergency response medical logistics is constructed. Real-world locations are used as shown in Figure 8, only to calibrate the network geometry, while the travel time is approximated by the straight-line distance divided by the drone's cruising speed. Representative drone platform and laboratory capacity parameters are grounded in publicly available specifications and throughput information [38,39], shown in Table 4.

Since the stylized instance remains small enough to be handled by the exact MILP solver, it is used to validate the integrated model behavior and to examine how the TTCT responds to key operational parameters via sensitivity analyses.

Table 4. Parameters and settings of the stylized real-case scenario.

Item	Value/description
Scenario	Metropolitan hub-and-spoke emergency response setting with 5 temporary collection points and 2 qualified testing institutions
Geometry	Real-world locations
Drone platform	Medical delivery drone platform
Drone speed	10 m/s
Payload and flight range	20 km with a 1 kg payload
Per-trip specimen capacity	Q = 30, based on 30 g per specimen, on average
Operational times	*take* = 1 min; *place* = 1 min; *rest* = 2 min.
Institution A testing rate	v^2 *(INSA)*= 0.8 specimens/min
Institution B testing rate	v^2 *(INSB)* = 0.5 specimens/min

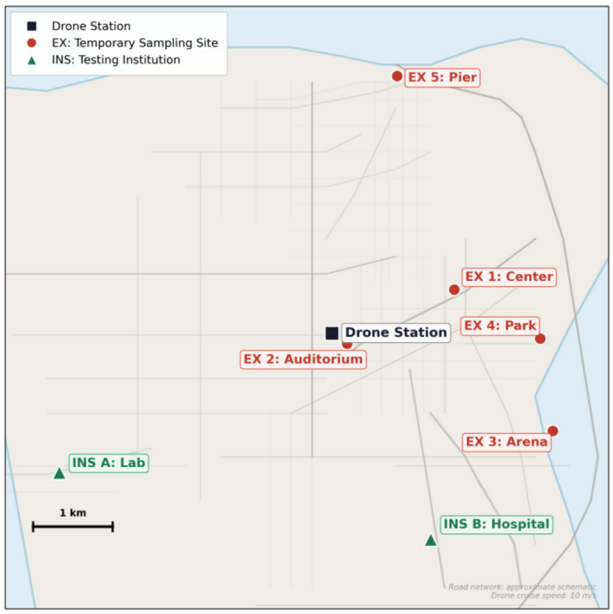

Figure 8. Locations of collection points and testing institutions in the stylized real-case scenario.

(1) Experimental results of the stylized real-case scenario

Table 5. Exact-solver results for the baseline stylized real-case scenario.

Scenario	Total_g	Exact_TTCT	oGap	CPUE(s)
SC1	1410	1096.54	1.09%	1200.00

Table 5 shows the exact-solver record for the stylized real-case scenario. *Total_g* denotes the total sample volume collected over all collection points, *Exact_TTCT* denotes the resulting total testing completion time, *oGap(%)* is the relative optimality gap at termination, and *CPUE(s)* is the solver runtime in seconds. In this baseline solution, the two testing institutions are almost perfectly balanced: *INS A* is busy for 1083.1 min with a 13.5 min idle time, whereas *INS B* is busy for 1087.1 min with a 9.5 min idle time.

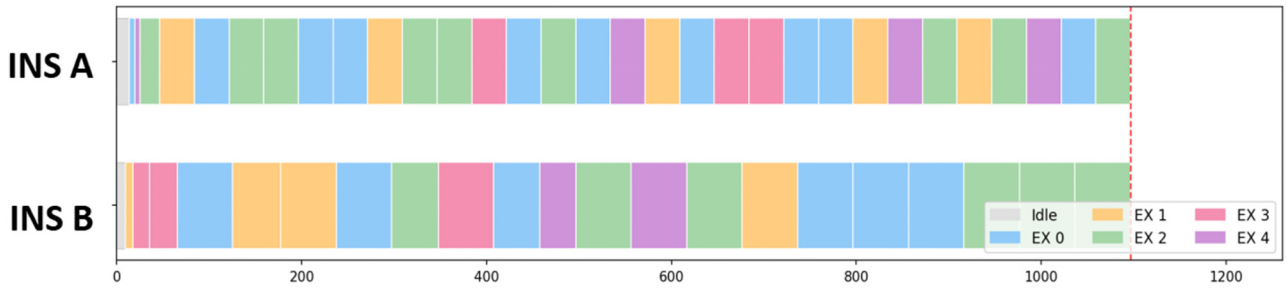

Figure 9. Workload distribution of the two testing institutions under the baseline solution.

The workload profile in Figure 9 shows how the integrated model uses repeated drone trips to route batches generated at different times to institutions at different distances, thereby synchronizing

delivery with downstream queueing by jointly shaping the temporal workload distribution across the two institutions.

(2) Sensitivity analysis

To further examine the validity and managerial implications of the model in Section 3, we conduct a set of exact-solver sensitivity tests around the baseline above. The tested scenarios vary the drone payload capacity Q, the common sampling rate V^1, the common testing rate V^2, the collection horizon L, the drone speed V^d, and selected node-specific perturbations in L, V^1, or V^2 at individual collection points and testing institutions, which are used to examine both scale effects and localized imbalance effects.

The columns *total_g, Exact_TTCT, oGap, CPUE(s)*, and *n_trips* denote the total sample volume, the resulting total testing completion time, the relative optimality gap, the solver runtime in seconds, and the number of drone trips, respectively. Cases where *CPUE(s)* reached the 1200-s time limit were stopped before the optimality gap fell below 1%. Nevertheless, all reported *oGap* remain small (at most 3.8%); thus, the solutions are sufficiently informative for sensitivity assessment. Overall, the sensitivity results are highly consistent with the model logic. Specifically, four insights emerge from Table 6.

Table 6. Exact-solver sensitivity analysis for the stylized real-case scenario.

Scenario	Setting	total_g	Exact_TTCT	oGap	CPUE(s)	n_trips
SC1	*Baseline*	1410	1096.54	1.09%	1200.00	52
SC2	*Q=10*	1410	1094.72	0.92%	976.72	147
SC3	*Q=180*	1410	1097.69	0.97%	139.14	15
SC4	$V^1 \times 0.5$	705	557.19	2.67%	1200.00	31
SC5	$V^1 \times 2.0$	2820	2189.52	0.93%	15.27	95
SC6	$V^2 \times 0.5$	1410	2190.88	0.99%	10.15	50
SC7	$V^2 \times 2.0$	1410	556.96	2.63%	1200.00	52
SC8	$L \times 0.5$	705	552.79	1.90%	1200.00	30
SC9	$L \times 2.0$	2820	2189.76	0.94%	94.88	98
SC10	V^1*(EX1)* $\times 0.2$	1050	818.82	1.36%	1199.99	41
SC11	V^1*(EX1)* $\times 5.0$	3210	2818.82	0.00%	0.83	108
SC12	V^1*(EX5)* $\times 0.2$	1294.8	1006.67	1.06%	1200.00	48
SC13	V^1*(EX5)* $\times 5.0$	1986	1542.47	0.96%	135.94	70
SC14	V^2*(INSA)* $\times 0.2$	1410	2155.69	0.90%	4.15	49
SC15	V^2*(INSA)* $\times 5.0$	1410	379.78	0.94%	13.02	49
SC16	V^2*(INSB)* $\times 0.2$	1410	1581.93	0.96%	73.84	53
SC17	V^2*(INSB)* $\times 5.0$	1410	444.15	3.80%	1200.00	52
SC18	*L(EX2)* $\times 2$	1842	1430.90	0.98%	20.40	66
SC19	*L(EX2)* $\times 5$	3138	2704.44	0.00%	0.34	105
SC20	$V^d \times 0.1$	1410	2028.26	0.96%	2.08	48
SC21	$V^d \times 0.2$	1410	1222.14	0.90%	331.33	51
SC22	$V^d \times 0.5$	1410	1102.12	1.59%	1200.00	53
SC23	$V^d \times 2$	1410	1095.30	0.98%	36.75	51

Payload capacity. Parameter Q is not a primary target for resource investment. Changing Q from 10 to 180 changes the TTCT only from 1094.72 to 1097.694, which is within 0.3% of the baseline, even though the required number of trips falls from 147 to 15. This pattern suggests that payload capacity mainly affects operational granularity and computational tractability rather than system-level completion performance.

Inspection capacity. The operational bottleneck lies in inspection capacity rather than transportation capacity. Relative to the baseline, halving the common testing rate nearly doubles the TTCT to 2190.88, whereas doubling it reduces the TTCT to 556.96. The node-specific cases are consistent: reducing one institution's testing rate to 20% of its baseline raises the TTCT to 2155.69, while increasing institutional capacity reduces the TTCT to 379.78 or 444.15.

Drone speed. In the stylized real-case scenario, drone speed is not a primary investment lever. Because *EX-INS* distances are short, moderate speed changes lead to little variation in the exact TTCT. The baseline case already shows that the two institutions remain almost perfectly balanced despite different delivery distances, and the final completion time is governed mainly by the downstream busy time rather than by transport differences. However, when *EX-INS* distances increase, or when the transportation speed is substantially reduced, speed can still become a critical bottleneck, which highlights the structural advantage of drones over ground vehicles.

Load imbalance. Load imbalance in L is more consequential than proportional changes in the total workload. When the workload becomes concentrated on a single origin, the exact TTCT increases markedly. For example, increasing the sampling rate at *EX1* fivefold increases the TTCT to 2818.82, which is substantially greater than the proportional system-wide increase $V^1 \times 2.0$ (SC5: 2189.52), even though the total sample volumes are of the same order. These results suggest that spatially uneven demand can amplify congestion beyond the effect of aggregate workload growth alone.

5.2. Generated instances: Heuristic performance and comparative analysis

To evaluate the heuristic framework on larger instances than the stylized real-case scenario, 20 collection points and 20 testing institutions are randomly generated within a rectangular area with a side length of 120. The time span for sample collection at each collection point is uniformly set to [0, 1000]. The sampling speed of the collection points and the testing speed of the testing institutions are assumed to follow a normal distribution with a mean of 0.8 and a variance of 0.2, being randomly generated. The specific values are shown in Table 7. Meanwhile, one drone for each collection point is initially assigned, with a speed of 1 and a capacity of 180.

The parameter values set in this study are based on the proportional relationships between these parameters in practice. Specifically, the settings for the drone's speed and capacity allow the farthest pair of collection points and testing institutions to satisfy $cir\_time \cdot v^1 > Q$, and the nearest pair of collection points and testing institutions to satisfy $cir\_time \cdot v^1 \le Q$. Thus, all scenarios considered in Section 4.1 are covered in the experiment.

Table 7. Operating speeds of collection points and testing institutions.

Collection point	Sampling speed	Testing institution	Testing speed
E1	0.64	I1	0.76
E2	0.68	I2	0.36
E3	0.51	I3	0.82
E4	0.65	I4	0.62
E5	0.70	I5	0.81
E6	0.88	I6	0.56
E7	0.72	I7	0.74
E8	1.11	I8	0.90
E9	0.92	I9	0.95
E10	0.52	I10	0.55
E11	1.05	I11	0.66
E12	0.54	I12	0.88
E13	0.87	I13	0.72
E14	0.87	I14	0.78
E15	0.99	I15	1.07
E16	0.71	I16	0.70
E17	0.59	I17	0.72
E18	0.69	I18	1.03
E19	0.78	I19	0.56
E20	0.66	I20	0.54

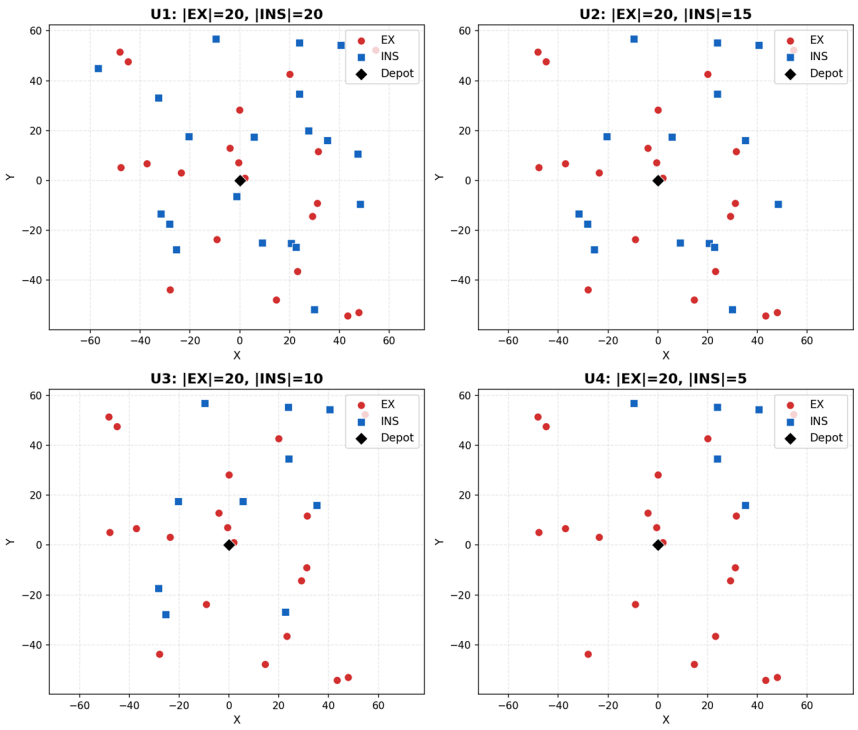

Figure 10. Distribution of collection points and testing institutions of instances U1–U4.

Using the base setting above, U1, U2, U3, and U4 vary the number of available testing institutions, which is set to 20, 15, 10, and 5, respectively, to represent increasing congestion at the testing stage. Instances U5–U8 shorten the collection timespan of half of the collection points from [0,1000] to [0,500]. Instances U9–U12 double the sampling speeds of half of the collection points. Instances U13–U16 double the testing speeds of half of the testing institutions. The representative spatial layouts for U1–U4 are shown in Figure 10.

(1) Experimental results for the generated instances

The heuristic results are reported in Table 8. *Inst.* indicates the instance name, and |*EX*| and |*INS*| indicate the numbers of collection points and available testing institutions, respectively. *TTCT* denotes the objective value returned by the proposed heuristic framework, and *CPU(s)* is the corresponding runtime in seconds. *TTCT_E* denotes the best feasible objective value found by Gurobi within the 3600-s time limit, and *CPUE(s)* is the corresponding Gurobi runtime. *GAPE* measures the relative difference between *TTCT* and *TTCT_E*. In addition, *Sd.* denotes the standard deviation of testing completion times across institutions, and *Cv.* is the corresponding coefficient of variation. Small *Cv.* values indicate a strong balancing of institutional completion times.

Table 8. Computational results for the U1–U16 synthetic benchmark instances.

Inst.	|EX|	|INS|	TTCT (min)	CPU(s)	TTCT_E (min)	CPUE(s)	GAPE	Sd.	Cv.
U1	20	20	1319.52	0.54	-	3600	-	57.77	4.53%
U2	20	15	1628.17	0.99	1652.76	3600	−1.51%	26.90	1.68%
U3	20	10	2179.22	1.34	4524.62	3600	−107.63%	12.86	0.59%
U4	20	5	4012.89	2.26	3992.46	3600	0.51%	26.83	0.67%
U5	20	20	1197.26	0.08	1793.54	3600	−49.80%	123.53	11.14%
U6	20	15	1289.66	0.92	1475.50	3600	−14.41%	29.56	2.51%
U7	20	10	1689.19	1.34	1681.27	3600	0.47%	35.28	2.32%
U8	20	5	2899.66	2.16	2925.52	3600	−0.89%	13.97	0.48%
U9	20	20	1739.69	0.54	15281.29	3600	−778.39%	133.84	8.21%
U10	20	15	2162.40	1.38	7743.52	3600	−258.10%	216.32	10.95%
U11	20	10	3124.13	2.16	12371.25	3600	−295.99%	39.10	1.26%
U12	20	5	5745.08	1.74	16493.48	3600	−187.09%	33.43	0.59%
U13	20	20	1165.57	0.52	2066.89	3600	−77.33%	100.12	9.02%
U14	20	15	1207.39	0.97	1631.63	3600	−35.14%	34.86	2.78%
U15	20	10	1566.77	0.91	1747.64	3600	−11.54%	67.33	4.06%
U16	20	5	2899.66	1.81	2797.75	3600	3.51%	10.52	0.35%

As shown in Table 8, Gurobi did not certify an optimal solution within 3600 s for any of the 16 benchmark instances; additionally, for instance U1, it did not find a feasible solution within the time limit. Among the remaining 15 instances for which Gurobi returned feasible solutions, the proposed heuristic obtained better results than Gurobi in 12 cases, accounting for 80.0%, and the average *GAPE* over these cases was −151.49%. In contrast, the heuristic was inferior to Gurobi in only 3 cases, accounting for 20.0%, and the average *GAPE* over these cases was only 1.50%. In terms of computational efficiency, the average runtime of the proposed heuristic was only 1.227 s, which is significantly shorter than the 3600-s time limit used by Gurobi.

These results indicate that obtaining exact solutions for this problem within a practical time budget is highly challenging and strongly instance-dependent, and that the truncated results of Gurobi at 3600 s are not stable. In contrast, the performance of the proposed heuristic is relatively stable in terms of both computational time and solution quality. In the scenarios of this study, the coefficient of variation of the testing completion times across institutions was maintained at a relatively low level. Among the 16 scenarios, the coefficient of variation was less than 5% in 13 scenarios. The results indicate that the algorithm proposed in this paper can effectively balance the testing completion times of different institutions, thereby achieving the goal of reducing the total testing completion time.

(2) Comparative experimental results

The comparative experiment in this subsection is designed as a controlled algorithmic comparison. By strictly holding constant the transport platform and all physical operating parameters (e.g., cruise speed and payload capacity), the comparator represents a conventional spatial-only nearest-institution assignment operating under identical physical constraints. This controlled setting ensures that any observed reduction in TTCT is driven purely by the algorithmic joint optimization of the dispatch timing, the trip frequency, and downstream queue synchronization.

Table 9 shows the comparative results. *TTCT (min)* denotes the TTCT under the full approach proposed in this study. *TTCT_C (min)* denotes the TTCT under the spatial-only nearest-institution comparator under the same operating setting. *TTCT_O (min)* denotes the TTCT obtained by the proposed framework without the refined spatial decision, that is, after the initialized spatial decision and restricted temporal decision only. *Gap_C=(TTCT-TTCT_C)/TTCT_C, Gap_O=(TTCT-TTCT_O)/TTCT_O,* and *Gap_OC=(TTCT_O-TTCT_C)/TTCT_C.* Negative values indicate TTCT reductions relative to the corresponding comparator.

Table 9. Comparative results against the spatial-only nearest-institution comparator under the same operating setting.

Inst.	|EX|	|INS|	TTCT (min)	TTCT_C (min)	TTCT_O (min)	Gap_C	Gap_O	Gap_OC
U1	20	20	1319.52	2099.10	1749.06	−37.14%	−24.56%	−16.68%
U2	20	15	1628.17	2571.75	1818.34	−36.69%	−10.46%	−29.30%
U3	20	10	2179.22	3291.95	2332.25	−33.80%	−6.56%	−29.15%
U4	20	5	4012.89	5069.25	4259.24	−20.84%	−5.78%	−15.98%
U5	20	20	1197.26	2099.10	1197.26	−42.96%	0.00%	−42.96%
U6	20	15	1289.66	2268.26	1360.65	−43.14%	−5.22%	−40.01%
U7	20	10	1689.19	2410.38	1876.78	−29.92%	−10.00%	−22.14%
U8	20	5	2899.66	4088.92	3236.14	−29.09%	−10.40%	−20.86%
U9	20	20	1739.69	4067.58	1821.66	−57.23%	−4.50%	−55.21%
U10	20	15	2162.40	5075.50	2210.88	−57.40%	−2.19%	−56.44%
U11	20	10	3124.13	6541.54	3165.97	−52.24%	−1.32%	−51.60%
U12	20	5	5745.08	7205.06	5748.03	−20.26%	−0.05%	−20.22%
U13	20	20	1165.57	1392.68	1240.16	−16.31%	−6.01%	−10.95%
U14	20	15	1207.39	2268.26	1327.94	−46.77%	−9.08%	−41.46%
U15	20	10	1566.77	2410.38	2300.21	−35.00%	−31.89%	−4.57%
U16	20	5	2899.66	5069.25	4272.07	−42.80%	−32.13%	−15.73%
AVG						−37.60%	−10.01%	−29.58%

As shown in Table 9, the *TTCT* is lower than *TTCT_C* in all 16 instances, with an average reduction of 37.60%. This finding demonstrates the value of TTCT-oriented integrated scheduling relative to the spatial-only nearest-institution comparator. *TTCT_O* is also lower than *TTCT_C* in all cases, with an average reduction of 29.58%, indicating that the initialized spatial decision already improves over pure distance-based allocation by incorporating workload balance. Moreover, *TTCT* is lower than *TTCT_O* by 10.01%, on average, confirming that the refined spatial decision further reduces the TTCT by flexibly redirecting trips to underutilized institutions.

The relative gain from integrated scheduling is particularly pronounced under a stronger imbalance. For U5–U8, where collection horizons are heterogeneous, the average *Gap_OC* is 31.50%, which is higher than the 22.78% for U1–U4. For U9–U12, where half of the collection points have doubled sampling speeds, the average *Gap_OC* increases to 45.85%, indicating that distance-only allocation performs poorly when high-generation sites create localized backlog. For U13–U16, where half of the testing institutions have doubled processing speeds, the average *Gap_O* increases to 19.78%, showing that flexible trip redirection becomes especially valuable when faster institutions create exploitable idle capacity.

Three key insights emerge from the comparative analysis.

From distance minimization to spatiotemporal synchronization: The consistent TTCT reduction over the nearest-institution comparator confirms that minimizing the delivery distance is a poor proxy for minimizing the turnaround time. Because testing institutions operate as downstream queuing systems, the scheduling objective must explicitly synchronize upstream dispatch with evolving downstream queue states rather than simply assigning each collection point to its closest institution.

Differentiated roles of the two spatial decisions under distinct imbalances: The initialized and refined spatial decisions address structurally different bottlenecks. Under a supply-side imbalance in terms of sampling speeds, workload-aware initialization contributes the dominant share of improvement by proactively dispersing surging workloads before queues form. Under capacity-side imbalance in terms of testing speeds, the refined spatial decision becomes critical, redirecting specific trips mid-operation to exploit fleeting idle capacity at faster institutions.

Computational viability for real-time dispatch: Commercial exact solvers fail to certify optimality for any benchmark instance within the time limit, whereas the proposed heuristic produces equal or better solutions in the vast majority of cases with runtimes on the order of one second. This combination of solution quality, workload balance, and computational speed is essential for time-sensitive medical logistics where dispatching decisions cannot await hour-long solver runs.

6. Conclusion

This paper studies the integrated scheduling problem of drone-based medical sample delivery and testing in hub-and-spoke networks. Departing from merely minimizing the in-transit delivery time, this work minimizes the system-wide TTCT by synchronizing multi-trip drone dispatching with continuous upstream sample collection and downstream deterministic FIFO testing. To capture the delivery–testing coupling mechanism underlying this problem, a MILP model is formulated, and an efficient three-stage heuristic framework is developed. This framework consists of the initialized spatial decision, restricted temporal decision, and refined spatial decision to decouple and optimize complex spatiotemporal variables.

The computational results validate both the soundness of the modeling logic and the practical merits of the proposed integrated scheduling strategy. In the stylized real-case scenario, the solutions obtained from the exact solver reveal that the integrated model can effectively balance downstream

workloads across testing institutions. Further sensitivity analyses indicate that testing capacity and workload imbalance are the primary drivers of the TTCT, while payload capacity and the drone speed exert comparatively minor impacts in compact networks. For larger benchmark instances, the proposed heuristic generates feasible solutions within approximately 1.23 seconds, on average, and reduces the system-wide TTCT by 37.60%, on average, compared with a spatial-only nearest-institution benchmark under identical operating conditions. Moreover, the two spatial decision stages play distinct roles: the workload-aware initialized spatial decision delivers outstanding benefits under sampling-side imbalance, while the refined spatial decision becomes more critical when heterogeneous testing capacities leave exploitable idle capacity.

These findings have clear managerial implications: for time-sensitive medical sample logistics, minimizing the delivery distance alone is not a sufficient substitute for optimizing overall turnaround performance. Instead, the spatiotemporal synchronization between upstream dispatching and downstream queueing states is the decisive factor. In particular, multi-trip scheduling achieves the greatest benefits when collection horizons or sampling rates are highly uneven or when testing institutions exhibit substantial discrepancies in their processing capacity. From a computational perspective, commercial exact solvers fail to certify optimality for the benchmark instances within the time limit because of the computational intractability of the MILP model, whereas the proposed heuristic maintains stable solution quality and a short runtime, supporting its deployment in practical real-time dispatch scenarios.

This study has limitations. The current model assumes deterministic collection and testing rates and simplifies battery replacement, regulatory restrictions, and environmental uncertainty into stylized operational constraints. Future research can extend the proposed framework to accommodate stochastic sample arrivals, drone malfunctions, weather volatility, richer regulatory constraints, heterogeneous drone fleets, multistation networks, and dynamic rolling-horizon decision environments.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (72301120, 72442009 and 72032001), Chenguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission (25CGA56), and the Fundamental Research Funds for the Central Universities.

Use of Generative-AI tools declaration

During the preparation of the final version of this manuscript, ChatGPT (OpenAI) was used to assist with language polishing, reference formatting, and document formatting. The authors reviewed and edited the output and take full responsibility for the content of this publication.

Author contributions

Qiuchen Gu: Conceptualization, methodology, model development, software, validation, formal analysis, visualization, writing-original draft. Jingyi Chen: Methodology, investigation, validation, writing-review & editing, and funding acquisition. Qi Zheng: Data curation, validation, and writing-review & editing. Bin Hu: Supervision, and writing-review & editing. Tijun Fan: Conceptualization,

supervision, project administration, funding acquisition, and writing-review & editing. All authors have read and approved the final manuscript.

Conflict of interest

All authors declare no conflicts of interest in this paper.

Data availability statement

The data that support the findings of this study are openly available in Zenodo at http://doi.org/10.5281/zenodo.17097362.

References

1. World Health Organization, Strengthening diagnostics capacity, World Health Organization, 2022. Available from: https://www.who.int/activities/strengthening-diagnostics-capacity.
2. M. Plebani, Clinical laboratory: Bigger is not always better, *Diagnosis*, **5** (2018), 41–46. https://doi.org/10.1515/dx-2018-0019
3. N. Stierlin, M. Risch, L. Risch, Current advancements in drone technology for medical sample transportation, *Logistics*, **8** (2024), 104. https://doi.org/10.3390/logistics8040104
4. M. P. Nisingizwe, P. Ndishimye, K. Swaibu, L. Nshimiyimana, P. Karame, V. Dushimiyimana, et al., Effect of unmanned aerial vehicle (drone) delivery on blood product delivery time and wastage in Rwanda: A retrospective, cross-sectional study and time series analysis, *Lancet Glob. Health*, **10** (2022), e564–e569. https://doi.org/10.1016/S2214-109X(22)00048-1
5. C. C. Murray, A. G. Chu, The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery, *Transp. Res. Part C Emerg. Technol.*, **54** (2015), 86–109. https://doi.org/10.1016/j.trc.2015.03.005
6. A. Otto, N. Agatz, J. F. Campbell, B. Golden, E. Pesch, Optimization approaches for civil applications of unmanned aerial vehicles (UAVs) or aerial drones: A survey, *Networks*, **72** (2018), 411–458. https://doi.org/10.1002/net.21818
7. J. Pasha, Z. Elmi, S. Purkayastha, A. M. Fathollahi-Fard, Y. E. Ge, Y. Y. Lau, et al., The drone scheduling problem: A systematic state-of-the-art review, *IEEE Trans. Intell. Transp. Syst.*, **23** (2022), 14224–14247. https://doi.org/10.1109/TITS.2022.3155072
8. S. H. Chung, B. Sah, J. Lee, Optimization for drone and drone-truck combined operations: A review of the state of the art and future directions, *Comput. Oper. Res.*, **123** (2020), 105004. https://doi.org/10.1016/j.cor.2020.105004
9. O. Dukkanci, J. F. Campbell, B. Y. Kara, Facility location decisions for drone delivery: A literature review, *Eur. J. Oper. Res.*, **316** (2024), 397–418. https://doi.org/10.1016/j.ejor.2023.10.036
10. K. Dorling, J. Heinrichs, G. G. Messier, S. Magierowski, Vehicle routing problems for drone delivery, *IEEE Trans. Syst. Man Cybern. Syst.*, **47** (2017), 70–85. https://doi.org/10.1109/TSMC.2016.2582745

11. C. Cheng, Y. Adulyasak, L. M. Rousseau, Drone routing with energy function: Formulation and exact algorithm, *Transp. Res. Part B Methodol.*, **139** (2020), 364–387. https://doi.org/10.1016/j.trb.2020.06.011
12. C. Cheng, Y. Adulyasak, L. M. Rousseau, Robust drone delivery with weather information, *Manuf. Serv. Oper. Manag.*, **26** (2024), 1402–1421. https://doi.org/10.1287/msom.2022.0339
13. Y. Shi, Y. Lin, B. Li, R. Li, A bi-objective optimization model for the medical supplies' simultaneous pickup and delivery with drones, *Comput. Ind. Eng.*, **171** (2022), 108389. https://doi.org/10.1016/j.cie.2022.108389
14. L. Zhang, Y. Ding, H. Lin, Optimizing synchronized truck-drone delivery with priority in disaster relief, *J. Ind. Manag. Optim.*, **19** (2023), 5143–5162. https://doi.org/10.3934/jimo.2022166
15. O. Dukkanci, A. Koberstein, B. Y. Kara, Drones for relief logistics under uncertainty after an earthquake, *Eur. J. Oper. Res.*, **310** (2023), 117–132. https://doi.org/10.1016/j.ejor.2023.02.038
16. A. Otto, B. Golden, C. Lorenz, Y. Luo, E. Pesch, L. A. Rocha, On delivery policies for a truck-and-drone tandem in disaster relief, *IISE Trans.*, **57** (2025), 1198–1214. https://doi.org/10.1080/24725854.2024.2410353
17. K. Zhang, Y. Shi, H. Guo, Y. Sun, Optimal decision-making for dispatching emergency supplies for natural disasters in mountainous areas based on truck-drone collaboration, *Chin. J. Manag. Sci.*, **33** (2025), 150–160. https://doi.org/10.16381/j.cnki.issn1003-207x.2023.1278
18. F. Wang, H. Li, H. Xiong, Truck-drone routing problem with stochastic demand, *Eur. J. Oper. Res.*, **322** (2025), 854–869. https://doi.org/10.1016/j.ejor.2024.11.036
19. X. Yang, Z. He, Y. Liu, S. Liu, Solving the vehicle-drone pickup and delivery problem in road congestion: A heuristic and its deep reinforcement learning-based improvement, *J. Ind. Manag. Optim.*, **21** (2025), 1630–1654. https://doi.org/10.3934/jimo.2024141
20. R. C. Hawkins, Laboratory turnaround time, *Clin. Biochem. Rev.*, **28** (2007), 179–194.
21. E. Yücel, F. S. Salman, E. S. Gel, E. L. Örmeci, A. Gel, Optimizing specimen collection for processing in clinical testing laboratories, *Eur. J. Oper. Res.*, **227** (2013), 503–514. https://doi.org/10.1016/j.ejor.2012.10.044
22. Z. B. Zabinsky, P. Dulyakupt, S. Zangeneh-Khamooshi, C. Xiao, P. Zhang, S. Kiatsupaibul, et al., Optimal collection of medical specimens and delivery to central laboratory, *Ann. Oper. Res.*, **287** (2020), 537–564. https://doi.org/10.1007/s10479-019-03260-9
23. S. J. Kim, G. J. Lim, J. Cho, M. J. Côté, Drone-aided healthcare services for patients with chronic diseases in rural areas, *J. Intell. Robot. Syst.*, **88** (2017), 163–180. https://doi.org/10.1007/s10846-017-0548-z
24. S. Chakraborty, R. A. Nadar, A. Tiwari, Designing a drone assisted sample collection and testing system during epidemic outbreaks, *J. Global Oper. Strategic Sourcing*, **15** (2022), 283–305. https://doi.org/10.1108/JGOSS-02-2021-0010
25. N. Stierlin, F. Loertscher, H. Renz, L. Risch, M. Risch, Preanalytic integrity of blood samples in uncrewed aerial vehicle (UAV) medical transport: A comparative study, *Drones*, **8** (2024), 517. https://doi.org/10.3390/drones8090517
26. S. Molinari, R. Patriarca, M. Ducci, The challenges of blood sample delivery via drones in urban environment: A feasibility study through specific operation risk assessment methodology, *Drones*, **8** (2024), 210. https://doi.org/10.3390/drones8050210

27. J. Zhao, D. Zhang, C. Li, P. Hu, A safety interval assessment method of medical drone transportation in urban low-altitude environment, *Ind. Eng. J.*, **28** (2025), 40–48. https://doi.org/10.3969/j.issn.1007-7375.240385
28. X. Huang, F. Ren, M. Liu, P. Jin, Y. Sun, Systematic research and application of a 5G medical unmanned aerial vehicle to deliver COVID-19 nucleic acid samples, *Health Secur.*, **22** (2024), 304–310. https://doi.org/10.1089/hs.2023.0090
29. Z. L. Chen, Integrated production and outbound distribution scheduling: Review and extensions, *Oper. Res.*, **58** (2010), 130–148. https://doi.org/10.1287/opre.1080.0688
30. S. Moons, K. Ramaekers, A. Caris, Y. Arda, Integrating production scheduling and vehicle routing decisions at the operational decision level: A review and discussion, *Comput. Ind. Eng.*, **104** (2017), 224–245. https://doi.org/10.1016/j.cie.2016.12.010
31. Y. Fu, Z. Li, K. Gao, A. M. Fathollahi-Fard, Z. Zhang, Integrated scheduling of distributed manufacturing with assembly and distribution: State of the art, challenges, and future directions, *Int. J. Prod. Res.*, (2025), 1–42. https://doi.org/10.1080/00207543.2025.2602898
32. J. Hurink, S. Knust, Tabu search algorithms for job-shop problems with a single transport robot, *Eur. J. Oper. Res.*, **162** (2005), 99–111. https://doi.org/10.1016/j.ejor.2003.10.034
33. C. A. Ullrich, Integrated machine scheduling and vehicle routing with time windows, *Eur. J. Oper. Res.*, **227** (2013), 152–165. https://doi.org/10.1016/j.ejor.2012.11.049
34. Y. Hou, Y. Fu, K. Gao, H. Zhang, A. Sadollah, Modelling and optimization of integrated distributed flow shop scheduling and distribution problems with time windows, *Expert Syst. Appl.*, **187** (2022), 115827. https://doi.org/10.1016/j.eswa.2021.115827
35. T. Liang, L. Zhou, Z. Jiang, Integrated scheduling of production and material delivery for the intelligent manufacturing system, *Int. J. Prod. Res.*, **63** (2025), 882–903. https://doi.org/10.1080/00207543.2024.2363435
36. Z. Zhou, E. Chen, R. Li, W. Sun, J. Shi, Research on emergency blood delivery problem based on multiple drones, *Chin. J. Manag. Sci.*, **33** (2025), 171–184. https://doi.org/10.16381/j.cnki.issn1003-207x.2023.1557
37. M. A. Lejeune, W. Ma, Drone-delivery network for opioid overdose: Nonlinear integer queueing-optimization models and methods, *Oper. Res.*, **73** (2025), 86–108. https://doi.org/10.1287/opre.2022.0489
38. Swiss Post, Matternet M2 specifications [PDF], 2017. Available from: https://www.post.ch/-/media/post/ueber-uns/medienmitteilungen/2017/drohnen/spezifikationen-matternet-m2.pdf
39. Roche Molecular Systems, Inc., cobas 6800/8800 systems: Systems specifications [PDF], 2015. Available from: https://diagnostics.roche.com/content/dam/diagnostics/Blueprint/en/pdf/rmd/RMD_cobas_6800_8000_07353898001_4_Systems_SpecTech_SS-1.pdf