Journal of Industrial and
Management Optimization

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

JIMO, 22(10): 5194–5226.
DOI:10.3934/jimo.2026179
Received: 01 July 2026
Revised: 14 September 2026
Accepted: 15 September 2026
Published: 20 September 2026

Research article

An LNS-based relax-and-restore matheuristic for integrated yard and manpower group planning in automotive terminals

Ziqiao Mei[1], **Feng Chen**[2,*]

[1] Department of Management Science & Engineering, Antai College of Economics & Management, Shanghai Jiao Tong University, 200030, Shanghai, China

[2] Sino-US Global Logistics Institute, Antai College of Economics & Management, Shanghai Jiao Tong University, 200030, Shanghai, China

* **Correspondence:** Email: fchen@sjtu.edu.cn.

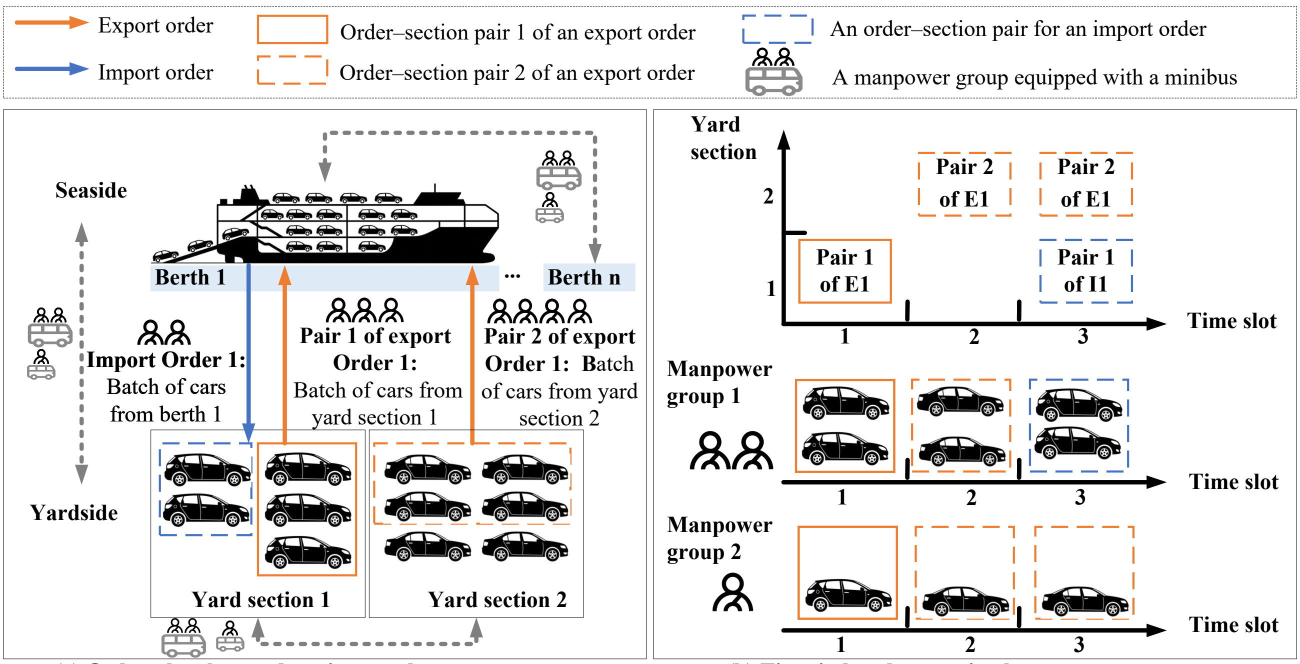
Abstract: This paper investigates a joint yard allocation and manpower group configuration and assignment problem in automotive roll-on/roll-off (RO-RO) terminals. Unlike existing yard allocation studies, this paper explicitly models manpower as standardized groups, each consisting of a fixed number of drivers transported together by a minibus between vessels and yard sections. Each order is associated with one vessel but can be allocated to multiple yard sections. The problem must determine the minimum required number of groups of each size as well as how to assign these groups to each order–section pair. We formulate the problem as an integer linear programming model to minimize the total manpower cost. To solve large-scale instances, we develop a large neighbourhood search-based relax-and-restore matheuristic (LNS-M) in which each encoded order sequence is evaluated through a three-stage relax-and-restore framework. The first stage constructs rough yard-operation plans, the second stage evaluates these plans with a partially relaxed manpower group configuration and assignment model, and the third stage conditionally restores integer feasibility for promising candidates. The main challenge introduced by group configuration is that extending the operation interval of an order–section pair can reduce the number of manpower groups assigned to this pair, but it may increase overlaps among several order–section pairs and create a higher peak driver demand. To address this challenge, we design a peak-aware scoring mechanism to determine the start and completion times of each pair. Extensive numerical experiments demonstrate the effectiveness of the proposed method and validate the contributions of the peak-aware scoring mechanism and the matheuristic framework. Compared with a rule-based heuristic, LNS-M reduces the average manpower cost by 21.35%.

Keywords: RO-RO terminal; yard allocation; group configuration; manpower assignment; matheuristic
Mathematics Subject Classification: 90C11, 90C59, 90C27, 90B06, 90B35

1. Introduction

Automotive terminals, also called roll-on/roll-off (RO-RO) terminals, form a critical interface in finished-car logistics, linking car manufacturers, inland distribution networks, and maritime transport services [1, 2]. Their operating performance affects not only vessel service reliability but also terminal operating cost. Unlike container terminals, which rely mainly on cranes, RO-RO terminals move cars directly between RO-RO vessels and yard sections. Loading and unloading operations are therefore tightly coupled with driver availability, internal transport distance, and yard occupation. This coupling makes manpower planning a central operational decision and manpower cost a key component of RO-RO terminal operating costs [3].

RO-RO terminal operations involve both import and export orders. An order represents a batch of cars associated with the same vessel and can be allocated to multiple yard sections. As shown in Figure 1 (a), for an import order, cars are unloaded from the vessel after berthing, parked in selected yard sections, and then leave the yard at the pickup time. For an export order, all cars arrive at the yard at the same time, are assigned to one or more yard sections, and are loaded onto the vessel before departure. Consequently, multiple order–section pairs generated by the same order may have their own planned start and completion times (see Figure 1 (b)), provided that these times remain within the order's feasible time window. These import and export processes jointly determine spatial and temporal yard occupation and the demand for drivers.



(a) Orders, berths, yard sections, and manpower groups (b) Time-indexed operation by manpower groups

Figure 1. Illustration of order operations by manpower groups in an automotive RO-RO terminal.

At a major RO-RO terminal in Shanghai, an operating day is divided into three shifts. At the start of each shift, the terminal decides how many drivers to employ from a third-party service provider and pays for all procured drivers for the entire shift. To facilitate operations, drivers are partitioned into standardized manpower groups, each with a fixed size and an assigned minibus that shuttles between yard sections and berths (see Figure 1 (a)). This group-based arrangement enhances operation coordination among drivers.

However, this group-based management also introduces a planning complexity. Because each group is indivisible, and each configured driver is paid for the entire shift, the choice of group sizes and quantities directly affects both the available drivers and the efficiency of yard operations. A larger group with more drivers per slot can lead to underutilized drivers when assigned to order–section pairs with small cars. A smaller group may mitigate such underutilization but increases the number of groups to coordinate and transport. The terminal thus faces a joint decision: not only where and when to allocate yard space but also how many groups of what sizes to configure and how to assign these indivisible groups to order–section pairs over time.

In current practice, these decisions are made sequentially. Planners first determine yard allocation and operation timing based on available yard capacity and experience and then assign manpower groups to operate the resulting order–section pairs. Because group assignment is independent of yard allocation, this sequential practice may overlook the feedback from group assignment to operation duration, yard occupation intervals, and peak manpower demand. The practical knowledge gap therefore lies in the lack of an integrated approach that jointly optimizes these interdependent decisions.

This paper investigates the joint optimization of yard allocation, manpower group configuration, and time-indexed group assignment in automotive RO-RO terminals. The joint consideration of these decisions is necessary because yard sections are heterogeneous, differing not only in capacity but also in their distance to various berths. Consequently, when assigned to a farther yard section, the same manpower group operates fewer cars within a given time slot. Therefore, the terminal must decide which yard sections to allocate to each order as well as the specific start and completion times for each order–section pair. These yard decisions determine the time-dependent storage occupation of the yard and subsequently affect the transport distance and operation efficiency of later cars. Simultaneously, because drivers are managed as standardized manpower groups, the terminal cannot assign an arbitrary, continuous number of individual drivers to orders. Instead, it must determine how many standardized manpower groups to configure within a shift and how to assign these groups to various order–section pairs over time. From the terminal operator's perspective, the objective of this paper is to minimize the total manpower cost while ensuring that all import and export operations are completed within their specified time windows and that yard capacity constraints are satisfied.

Solving this joint optimization problem is challenging. Beyond the strong coupling of spatial and temporal variables mentioned above, group configuration creates a stretch–compress trade-off. Extending the operation interval of an order–section pair can reduce the number of groups required for that pair. However, extending several operations simultaneously may increase their temporal overlap and raise the peak number of simultaneously required groups, thereby forcing additional complete groups to be configured and paid for the entire shift. Conversely, compressing selected operations may reduce temporal overlap but require more groups within fewer time slots.

Figure 2 illustrates this trade-off with seven order–section pairs and six time slots. Pair A has six group-slots of work, and each of the other six pairs has four group-slots. These workloads are fixed across the three plans. In the extremely compressed plan in Figure 2(a), each pair is processed in one slot, so the peak reaches eight groups in Figure 2(d). In the moderately staggered plan in Figure 2(b), Pair A spans all six slots, whereas the other pairs are staggered in two-slot intervals. This plan creates a flat profile of five groups in Figure 2(e), which reaches the average workload lower bound. In the extremely stretched plan in Figure 2(c), every pair uses only one group, and Pair A spans all six slots, but the other stretched pairs overlap heavily in the middle of the horizon, so the peak rises to seven groups in Figure 2(f).

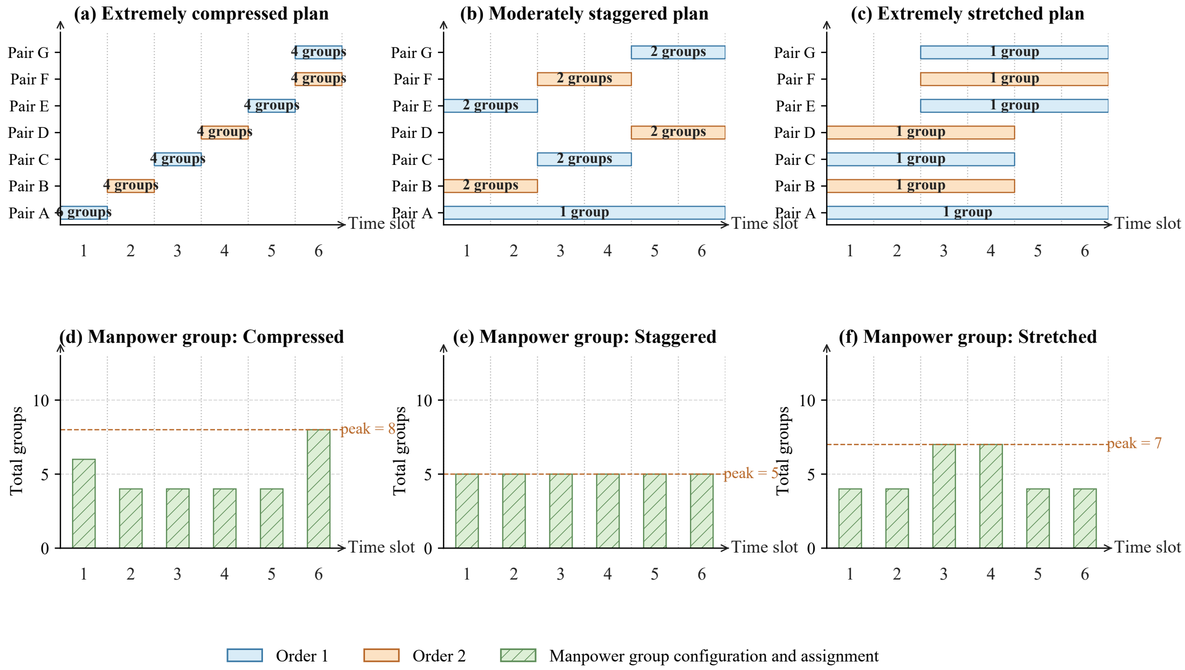


Figure 2. Illustration of the stretch–compress trade-off.

The main contributions of this paper are as follows.

1. We formulate an integer linear programming model that integrates yard allocation, manpower group configuration, and manpower group assignment. The model treats standardized manpower groups as indivisible dispatching resources and jointly optimizes yard-section allocation, operation timing, group configuration, and time-indexed group assignment.

2. We identify a group-configuration-induced stretch–compress trade-off. Extending an order–section operation may reduce its assigned group requirement, but overlapping extended operations can increase the peak number of configured groups. To address this mechanism, we design a peak-aware scoring rule that selects start and completion times by considering both local operation efficiency and system-level peak group demand.

3. We develop a large neighbourhood search-based relax-and-restore matheuristic (LNS-M) for large-scale instances. The outer LNS searches over order sequences, and an inner three-stage framework constructs rough yard-operation plans, screens them through a partially relaxed group-configuration-and-assignment model, and conditionally restores integer-feasible assignments for promising candidates.

4. We demonstrate the effectiveness of the proposed method through extensive numerical experiments. LNS-M matches CPLEX optimal solutions on tiny instances, obtains reliable feasible solutions on

larger instances, and reduces the average manpower cost by 21.35% compared with a rule-based heuristic. The sensitivity analysis further suggests that maintaining a diverse set of group sizes rather than simply increasing the number of identical groups can preserve configuration flexibility and avoid costly over-configuration.

The remainder of the paper is organized as follows. Section 2 reviews related studies on RO-RO terminal operations, yard planning, and manpower scheduling. Section 3 describes the integrated planning problem and presents the integer formulation. Section 4 develops the LNS-based relax-and-restore matheuristic. Section 5 reports the computational study design and results. Section 6 concludes the paper and discusses future research directions.

2. Literature review

We first examine RO-RO yard allocation, manpower planning, and integrated terminal operations to identify the research gap in RO-RO terminals. We then look beyond the RO-RO literature to configurable workforce and resource scheduling, as well as related integrated manpower planning, to explore potential solution ideas for resource configuration and joint optimization.

RO-RO yard allocation. Cordeau et al. [2] assigned car groups to parking rows in an automotive trans-shipment terminal to minimize handling time while considering managerial rules such as adjacency and no relocation. Chen et al. [4] optimized storage location assignment in an automotive RO-RO terminal by using car-group compactness, attraction degree, and a rolling-horizon heuristic to improve ship-loading efficiency. Zhang and Fan [5] proposed a resilient yard template with shared storage lanes for RO-RO ports and optimized transportation distance, space utilization, and operation interference. Although yard optimization is well-developed, it overlooks joint manpower optimization: Every yard allocation plan must ultimately be executed by specific manpower resources.

Manpower and integrated planning in RO-RO terminals. Mattfeld and Kopfer [1] developed a planning system that links manpower planning and inventory control in car trans-shipment through hierarchical decomposition and heuristics. Mei et al. [6] studied a manpower-dependent berth allocation problem in which manpower capacity affects vessel operation time and berth decisions. Jaehn et al. [7] formulated a constraint-programming model for multimode, multiskill personnel scheduling in RO-RO loading. Although these three studies incorporate manpower into terminal planning, none jointly determines yard allocation. Recently, Chen et al. [8] further incorporated shuttle-bus stop-point decisions into the joint optimization of storage-location assignment and driver-task allocation, aiming to improve work efficiency and fairness among drivers. However, their model assumes a fixed driver pool and treats drivers as individually assignable manpower units. As such, it does not address how many standardized manpower groups of each size should be configured. Overall, the study by Zhang et al. [9] is the closest to ours. They formulated a mixed-integer model that jointly optimizes time-indexed yard allocation and driver assignment and solved it using a column-generation-based heuristic algorithm. Their formulation dynamically tracks car accumulation and release, whereas ours adopts a conservative capacity-reservation mechanism. More importantly, they treated driver capacity as exogenous and minimize car driving distance, so peak driver demand affects feasibility but not cost as long as the capacity limit is satisfied. In our model, manpower-group configuration is endogenous and paid for the entire shift, so the peak requirement directly affects manpower cost. This difference changes both the problem structure and the required solution approach. Table 1 provides a detailed comparison.

Table 1. Focused comparison with Zhang et al. [9].

Study	Yard allocation	Dynamic car accumulation and release	Time-indexed manpower assignment	Assignment for number of drivers	Standardized indivisible manpower-group configuration	Standardized indivisible manpower-group assignment
Zhang et al. [9]	✓	✓	✓	✓	–	–
This paper	✓	–	✓	–	✓	✓

Table 2. Positioning among manpower-integrated planning studies in RO-RO terminals.

Study	Decisions on manpower	Main coupling	Objective	Algorithm
Mattfeld and Kopfer [1]	Shift-level driver demand; task-to-gang assignment and sequencing	Driver demand–task timing–storage location	Minimize storage cost and deviations from regular driver levels	Hierarchical decomposition and heuristics
Zhang et al. [9]	Number of drivers assigned to each order–section task in each time slot	Driver assignment–dynamic yard occupation	Minimize total car driving distance	Rule-based stepwise algorithm, reinforced column generation, and multiroot branch-and-bound
Mei et al. [6]	Manpower-ability units assigned to each vessel in each time slot	Manpower ability–operation time–berth congestion	Minimize total vessel turnaround time	Hybrid adaptive large neighborhood search; column generation for lower bounds
Jaehn et al. [7]	Internal and external multiskilled teams assigned to batch tasks and operating modes	Team skills and modes–batch loading	Maximize loaded-batch priority; minimize external teams	Constraint programming
Chen et al. [8]	Driver-round composition and shuttle-bus stop-point decisions for each loading round	Driver rounds–shuttle stops–work time and fairness	Improve driver efficiency and workload fairness	Particle-swarm and adaptive-large-neighborhood-search metaheuristic; dynamic programming
This paper	**Number of groups by size and time-indexed group assignment to order–section tasks**	**Group configuration–duration–yard occupation**	**Minimize total manpower cost**	**LNS-based relax-and-restore matheuristic**

Beyond the RO-RO-specific literature, two broader streams inform our work. **Configurable workforce and resource scheduling:** Li et al. [10] dynamically allocate individual workers under uncertain orders. Tian et al. [11] formulated assembly scheduling with flexible workforce configuration. Wen et al. [12] scheduled airline crew members individually. Mandal et al. [13] jointly optimized workforce size and shift schedules for last-mile delivery. Calik et al. [14] jointly determined a heterogeneous electric fleet and its routes, resembling discrete group configuration. They demonstrated that configuration decisions carry intrinsic value and should be optimized endogenously rather than being fixed a priori. **Related integrated resource planning:** Container-terminal studies integrate berth, crane, truck, and yard decisions [15–20]. Wang et al. [21] studied the integration of value-added services and automobile distribution in RO-RO terminals. Joint optimization models have also been developed by Ko [22] for airline pricing and seat allocation, Tang and Chu [23] for unit-load-device dispatch and scale, and Tang and Wang [24] for transportation outsourcing under stochastic demand. These works show that resource decisions must be coupled with spatial and temporal operations. However, these existing frameworks cannot be directly adapted to our problem, as they do not account for the distinctive RO-RO terminal characteristics. In our setting, the endogenous configuration of standardized manpower groups simultaneously influences the operation duration of each order–section pair and the resulting yard allocation.

In summary, the research gap lies in the RO-RO-specific coupling of yard occupation, order time windows, operation duration, and manpower-group configuration and assignment. Table 2 summarizes the differences among manpower-related studies in RO-RO terminals. The external literature confirms the value of workforce configuration and integrated resource planning but fails to capture the feedback central to our problem, motivating the development of a dedicated joint formulation and solution.

3. Problem description and formulation

3.1. Problem description

Given an automotive RO-RO terminal serving a set of car orders, let $\mathcal{R}$ denote the set of all orders. The set is partitioned into import orders $\mathcal{R}^I$, export orders with loading operations in the current shift $\mathcal{R}^E$, and export orders with fixed yard occupation but no loading operation in the current shift $\mathcal{R}^S$; that is $\mathcal{R} = \mathcal{R}^I \cup \mathcal{R}^E \cup \mathcal{R}^S$. The yard area is divided into a set of yard sections $\mathcal{I}$, and the planning horizon of one working shift is discretized into a set of time slots $\mathcal{T} = \{0, 1, \ldots, T-1\}$. For each yard section $i \in \mathcal{I}$ and time slot $t \in \mathcal{T}$, the available parking capacity Y_{it} is known in advance. The quayside contains several berth sections. For each order $r \in \mathcal{R}$, its target berth is known, and the operational efficiency between this berth and each yard section i is represented by β_{ri}. The value of β_{ri} is pre-estimated from the operating route. Let S_k be the number of drivers in one manpower group of type $k \in \mathcal{K}$, and let D_k be the maximum number of type-k manpower groups that can be configured in the shift.

Each order r containing N_r cars is operated by manpower groups within a known feasible time window $[L_r, U_r]$, and its yard occupation is tracked over the planning horizon $[0, T]$. To ensure operational feasibility in real-world operations, we adopt a capacity reservation strategy for yard allocation rather than dynamically tracking the incremental accumulation of yard occupation.

For an import order $r \in \mathcal{R}^I$, cars are unloaded from the vessel to selected yard sections, occupying yard capacity until a predetermined pickup time b_r. These import orders may either complete their operations within the current shift, as shown in Figure 3 (a), or continue their yard occupation beyond the shift boundary (Figure 3 (b)). Conversely, for an export order $r \in \mathcal{R}^E$, cars arrive at the yard at an effective time a_r, occupy selected sections, and are driven to the target berth during the loading operation. Such export orders may arrive and load entirely within the current shift, as shown in Figure 3 (c), or may have been stored before the shift and be loaded during it (Figure 3 (d)). Finally, export orders in $\mathcal{R}^S$ arrive at and occupy the yard during the current shift, but their vessels do not berth within the shift, and thus, no loading operation is performed (Figure 3 (e)). These orders consume yard capacity without requiring any manpower group assignment, meaning that $z_{ritk} = 0$ across all time slots and group types.

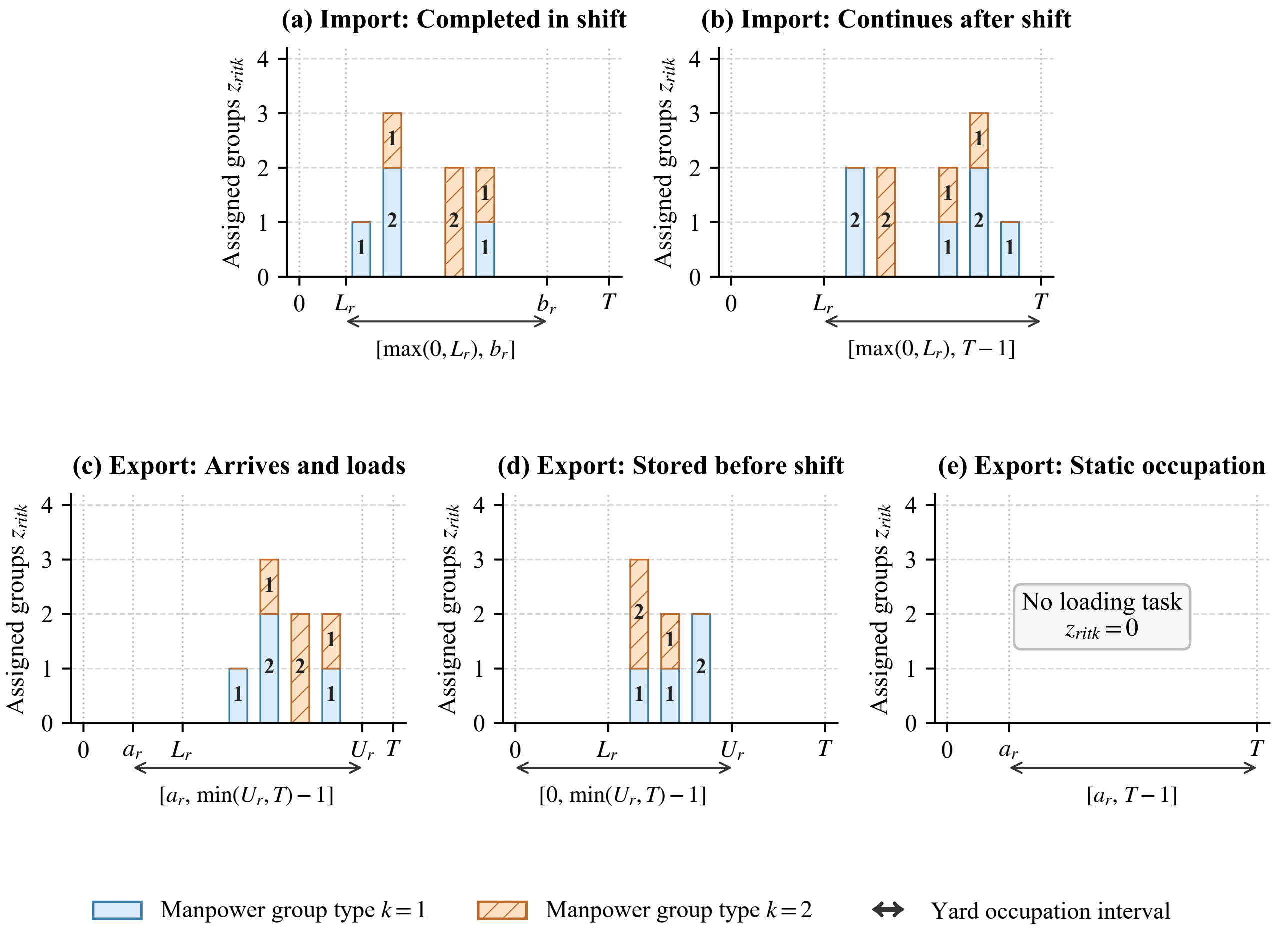


Figure 3. Operation timelines and time-indexed manpower group assignment scenarios.

For the terminal operator, three related plans need to be made. The first is a yard-section allocation plan, which determines the base quantity x_{ri} of order r assigned to yard section i. The second determines the planned start and completion times of each allocated order–section pair (r, i). The third is a manpower group configuration and assignment plan, which determines the number g_k of configured manpower groups of each type k and the number z_{ritk} of type-k manpower groups assigned to order r in yard section i at time slot t.

Overall, the problem studied in this paper is to jointly determine yard allocation plans, manpower group configurations, and time-indexed group assignments for all orders. The objective is to minimize the manpower cost while ensuring that all import and export orders can be completed within their feasible time windows and that the time-dependent capacity of each yard section is not violated.

The model uses the following assumptions.

1. Planning is conducted for one shift.

2. Manpower groups are standardized according to minibus size and are dispatched as indivisible operational units.

3. Operational efficiencies between berths and yard sections are known and deterministic within the planning horizon. An allowance for the transition time incurred when a manpower group switches between tasks is incorporated into the effective processing time used to estimate β_{ri}.

4. Operations for each order must be performed continuously between its actual start and finish times, and preemption is not allowed once an operation starts.

3.2. Notation

Table 3 summarizes the notation used in the model.

Table 3. Notations.

Notation	Meaning
Sets	
$\mathcal{R}$	Set of all orders
$\mathcal{R}^I$	Set of import orders
$\mathcal{R}^E$	Set of export orders with loading or yard-side operations in the shift
$\mathcal{R}^S$	Set of export orders with fixed yard occupation and no loading operation in the shift
$\mathcal{I}$	Set of yard sections
$\mathcal{T} = \{0, 1, \ldots, T-1\}$	Set of time slots
$\mathcal{K}$	Set of standardized manpower group types
Parameters	
N_r	Number of cars of order r that must be assigned or operated
Y_{it}	Available capacity of yard section i at time t
β_{ri}	Number of cars that one driver can transfer per time slot for order r and yard section i
S_k	Number of drivers in one manpower group of type k
D_k	Maximum number of manpower groups of type k that can be configured
C	Manpower cost per driver per time slot
a_r	Effective arrival time of export order r in the yard
b_r	Bill-taking or pickup time of import order r
L_r	Earliest feasible planned start time of order r
U_r	Latest feasible planned completion time of order r
$\bar{y}_{ri}$	Known initial yard quantity of export order r in yard section i, if applicable
M	A sufficiently large positive constant
Decision variables	
x_{ri}	Base quantity of order r assigned to yard section i; for an export order with $a_r < 0$, this variable is fixed at $\bar{y}_{ri}$
y_{rit}	Number of cars of order r occupying yard section i at time t
z_{ritk}	Number of manpower groups of type k assigned to order r in yard section i at time t
g_k	Number of configured manpower groups of type k
u_{ri}	Binary variable equal to 1 if order r is allocated to yard section i
s_{rit}	Binary variable equal to 1 if order–section pair (r, i) starts at time t
e_{rit}	Binary variable equal to 1 if order–section pair (r, i) completes at time t

3.3. Integer linear programming model

Based on the above definitions, the integer linear programming (ILP) is formulated as follows:

$$\min \quad \sum_{k\in\mathcal{K}} CS_k T g_k. \tag{3.1}$$

The objective function (3.1) minimizes the manpower cost. For each group type k, $S_k g_k$ is the number of drivers configured for the shift, and multiplying this quantity by the shift length T and the manpower cost per driver per time slot C gives the corresponding full-shift manpower cost.

Order quantity and initial yard allocation. All import and export orders must be appropriately assigned to yard sections. These constraints are formulated as follows:

$$\sum_{i\in\mathcal{I}} x_{ri} = N_r, \qquad \forall r \in \mathcal{R}^I, \tag{3.2}$$

$$\sum_{i\in\mathcal{I}} x_{ri} = N_r, \qquad \forall r \in \mathcal{R}^E \cup \mathcal{R}^S \text{ with } a_r \geq 0. \tag{3.3}$$

Constraint (3.2) requires every import order to be assigned to yard sections. Constraint (3.3) requires every export order that arrives at the yard within the planning horizon to be assigned to yard sections. For each export order with $a_r < 0$, x_{ri} is fixed at the known value $\bar{y}_{ri}$ for all $i \in \mathcal{I}$.

Yard occupation evolution. For import orders, yard occupation evolves from the planned start time until pickup as follows:

$$y_{rit} \geq x_{ri} - M\left(1 - \sum_{\tau=0}^{t} s_{ri\tau}\right), \qquad \forall r \in \mathcal{R}^I,\ i \in \mathcal{I},\ t \leq b_r, \tag{3.4}$$

$$y_{rit} \leq x_{ri}, \qquad \forall r \in \mathcal{R}^I,\ i \in \mathcal{I},\ t \leq b_r, \tag{3.5}$$

$$y_{rit} \leq M \sum_{\tau=0}^{t} s_{ri\tau}, \qquad \forall r \in \mathcal{R}^I,\ i \in \mathcal{I},\ t \leq b_r, \tag{3.6}$$

$$y_{rit} = 0, \qquad \forall r \in \mathcal{R}^I,\ i \in \mathcal{I},\ t > b_r. \tag{3.7}$$

Constraint (3.4) ensures that the assigned quantity is reserved once the operation starts. Constraint (3.5) caps the occupation by the assigned quantity, and Constraint (3.6) forces occupation to zero before the planned start. Constraint (3.7) removes the occupation after pickup.

For export orders with loading operations, yard occupation is defined from arrival until completion as follows:

$$y_{rit} = 0, \qquad \forall r \in \mathcal{R}^E,\ i \in \mathcal{I},\ t < a_r, \tag{3.8}$$

$$y_{rit} \geq x_{ri} - M\left(1 - \sum_{\tau=t}^{T-1} e_{ri\tau}\right), \qquad \forall r \in \mathcal{R}^E,\ i \in \mathcal{I},\ t \geq a_r, \tag{3.9}$$

$$y_{rit} \leq x_{ri}, \qquad \forall r \in \mathcal{R}^E,\ i \in \mathcal{I},\ t \geq a_r, \tag{3.10}$$

$$y_{rit} \leq M \sum_{\tau=t}^{T-1} e_{ri\tau}, \qquad \forall r \in \mathcal{R}^E,\ i \in \mathcal{I},\ t \geq a_r. \tag{3.11}$$

Constraint (3.8) sets yard occupation to zero before arrival. Constraint (3.9) reserves the assigned quantity after arrival while completion has not occurred. Constraint (3.10) caps occupation by the assigned quantity, and Constraint (3.11) removes occupation after completion.

For export orders without loading operations in the current shift, yard occupation and manpower assignment are specified as follows:

$$y_{rit} = 0, \qquad \forall r \in \mathcal{R}^S,\ i \in \mathcal{I},\ t < a_r, \tag{3.12}$$

$$y_{rit} = x_{ri}, \qquad \forall r \in \mathcal{R}^S,\ i \in \mathcal{I},\ t \geq a_r, \tag{3.13}$$

$$z_{ritk} = 0, \qquad \forall r \in \mathcal{R}^S,\ i \in \mathcal{I},\ t \in \mathcal{T},\ k \in \mathcal{K}. \tag{3.14}$$

Constraint (3.12) sets occupation to zero before arrival. Constraint (3.13) maintains the assigned quantity after arrival. Constraint (3.14) prohibits manpower-group assignment for these orders throughout the shift.

Yard capacity and manpower group availability. Yard capacity and manpower-group availability are constrained in each time slot as follows:

$$\sum_{r \in \mathcal{R}} y_{rit} \leq Y_{it}, \qquad \forall i \in \mathcal{I},\ t \in \mathcal{T}, \tag{3.15}$$

$$\sum_{r \in \mathcal{R}} \sum_{i \in \mathcal{I}} z_{ritk} \leq g_k, \qquad \forall k \in \mathcal{K},\ t \in \mathcal{T}, \tag{3.16}$$

$$g_k \leq D_k, \qquad \forall k \in \mathcal{K}. \tag{3.17}$$

Constraint (3.15) limits yard occupation in each time slot. Constraint (3.16) ensures that the number of type-k groups assigned in any time slot does not exceed the configured number of type-k groups. Constraint (3.17) limits the configured number of type-k groups by their availability.

Operation capacity. For each order–section pair, the assigned manpower groups must satisfy the following constraint:

$$\sum_{t \in \mathcal{T}} \sum_{k \in \mathcal{K}} z_{ritk} S_k \beta_{ri} \geq x_{ri}, \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I}. \tag{3.18}$$

Constraint (3.18) requires the total number of cars transferred by the assigned manpower groups over the planning horizon to cover the quantity assigned to each order–section pair.

Planned start and completion of order–section pairs. The time-related requirements for each allocated order–section pair are formulated as follows:

$$x_{ri} \leq M u_{ri}, \qquad \forall r \in \mathcal{R},\ i \in \mathcal{I}, \tag{3.19}$$

$$x_{ri} \geq u_{ri}, \qquad \forall r \in \mathcal{R},\ i \in \mathcal{I}, \tag{3.20}$$

$$\sum_{t \in \mathcal{T}} s_{rit} = u_{ri}, \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I}, \tag{3.21}$$

$$\sum_{t \in \mathcal{T}} e_{rit} = u_{ri}, \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I}, \tag{3.22}$$

$$\sum_{t \in \mathcal{T}} t s_{rit} \geq L_r - M(1 - u_{ri}), \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I}, \tag{3.23}$$

$$\sum_{t \in \mathcal{T}} t e_{rit} \geq \sum_{t \in \mathcal{T}} t s_{rit} - M(1 - u_{ri}), \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I}, \tag{3.24}$$

$$\sum_{t \in \mathcal{T}} t e_{rit} \leq U_r - 1 + M(1 - u_{ri}), \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I}. \tag{3.25}$$

Constraint (3.19) activates u_{ri} only when a positive quantity is assigned, whereas Constraint (3.20) prevents a yard allocation without an assigned quantity. Constraint (3.21) requires exactly one planned start time for every allocated order–section pair, and Constraint (3.22) requires exactly one planned completion time. Constraint (3.23) imposes the earliest feasible start time. Constraint (3.24) ensures that completion does not precede the start, and Constraint (3.25) imposes the latest feasible completion time.

Manpower-group assignment is restricted to the operation interval and must remain continuous as follows:

$$z_{ritk} \leq D_k \sum_{\tau=0}^{t} s_{ri\tau}, \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I},\ t \in \mathcal{T},\ k \in \mathcal{K}, \tag{3.26}$$

$$z_{ritk} \leq D_k \left(1 - \sum_{\tau=0}^{t-1} e_{ri\tau} \right), \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I},\ t \in \mathcal{T},\ k \in \mathcal{K}, \tag{3.27}$$

$$\sum_{k \in \mathcal{K}} z_{ritk} \geq u_{ri} - M \left(2 - \sum_{\tau=0}^{t} s_{ri\tau} - \sum_{\tau=t}^{T-1} e_{ri\tau} \right), \qquad \forall r \in \mathcal{R}^I \cup \mathcal{R}^E,\ i \in \mathcal{I},\ t \in \mathcal{T}. \tag{3.28}$$

Constraint (3.26) prohibits group assignment before the planned start time. Constraint (3.27) prohibits group assignment after the planned completion time. Constraint (3.28) requires at least one manpower group to operate in every time slot between the planned start and completion times of an allocated order–section pair.

Variable domains. The domains of all decision variables are specified by Constraints (3.29)–(3.32):

$$x_{ri}, y_{rit} \in \mathbb{Z}_+, \qquad \forall r \in \mathcal{R},\ i \in \mathcal{I},\ t \in \mathcal{T}, \tag{3.29}$$

$$z_{ritk} \in \mathbb{Z}_+, \qquad \forall r \in \mathcal{R},\ i \in \mathcal{I},\ t \in \mathcal{T},\ k \in \mathcal{K}, \tag{3.30}$$

$$g_k \in \mathbb{Z}_+, \qquad \forall k \in \mathcal{K}, \tag{3.31}$$

$$u_{ri}, s_{rit}, e_{rit} \in \{0, 1\}, \qquad \forall r \in \mathcal{R},\ i \in \mathcal{I},\ t \in \mathcal{T}. \tag{3.32}$$

4. Solution method

4.1. Overall algorithm framework

As the instance size increases, directly solving the ILP becomes computationally difficult. This is because it jointly determines yard-section allocation, time-window-constrained operation timing, manpower group-size configuration, and time-indexed group assignment under tightly coupled yard and manpower capacity constraints. These decisions create a large integer search space, and the group-configuration decisions further couple resource availability with operation duration and yard occupation. Because matheuristics have been successfully applied to complex resource-constrained allocation and scheduling problems [11, 25–27], we design an LNS-based relax-and-restore matheuristic (LNS-M). The outer LNS searches over order sequences, and the inner three-stage relax-and-restore evaluation framework constructs rough plans, screens them by a partially relaxed model, and selectively restores integer feasibility for promising candidates.

Throughout the solution method, symbol $\mathbf{x}$ denotes the complete solution vector, which collects all decision variables, including x_{ri}, y_{rit}, z_{ritk}, g_k, u_{ri}, s_{rit}, and e_{rit}.

The overall procedure is summarized in Algorithm 1.

Algorithm 1. Overall framework of the LNS-based relax-and-restore matheuristic.

Require: Instance data, number of independent runs N_{run}, number of starts S, iteration limit N, initial temperature $Temp_0$, minimum temperature $Temp_{\min}$, cooling factor α
Ensure: Best restored feasible solution $\mathbf{x}^*$

$\mathbf{x}^* \leftarrow \emptyset$
for $s = 1, \ldots, S$ **do**
  for $m = 1, \ldots, N_{\text{run}}$ **do**
    $(\pi^c, \check{\mathbf{x}}^c, best, UB) \leftarrow$ InitializeSequence(s, m)
    if $best$ is infeasible **then**
      continue
    end if
    $Temp \leftarrow Temp_0$
    for $n = 1, \ldots, N$ **do**
      Randomly remove q orders from π^c and store the removed set $\mathcal{R}^{rem}$
      Randomly reinsert the removed orders and obtain π'
      $(\check{\mathbf{x}}, V(\check{\mathbf{x}})) \leftarrow$ Evaluate(π')
      if $\check{\mathbf{x}}$ is feasible **then**
        if $V(\check{\mathbf{x}}) < UB$ **then**
          $\mathbf{x}^I \leftarrow$ Restore$(\check{\mathbf{x}})$
          if $\mathbf{x}^I$ is feasible and $Obj(\mathbf{x}^I) < UB$ **then**
            $best \leftarrow \mathbf{x}^I$, $UB \leftarrow Obj(\mathbf{x}^I)$
          end if
        end if
        if $V(\check{\mathbf{x}}) < V(\check{\mathbf{x}}^c)$ **or** AcceptBySA$(V(\check{\mathbf{x}}), V(\check{\mathbf{x}}^c), Temp)$ **then**
          $(\pi^c, \check{\mathbf{x}}^c) \leftarrow (\pi', \check{\mathbf{x}})$
        end if
      end if
      $Temp \leftarrow \max(Temp_{\min}, Temp \cdot \alpha)$
    end for
    if $\mathbf{x}^*$ is empty or $Obj(best) < Obj(\mathbf{x}^*)$ **then**
      $\mathbf{x}^* \leftarrow best$
    end if
  end for
end for
return $\mathbf{x}^*$

LNS-M starts from a multi-start structure, where each start is repeated through N_{run} independent runs to improve global exploration. For each start and run, InitializeSequence generates an initial order sequence and obtains an initial relaxed solution. The order sequence is the search representation used by the outer LNS. In each iteration, random removal and random insertion perturb the current sequence and generate a large-neighborhood candidate sequence. The candidate is then evaluated by the three-stage relax-and-restore framework. Stage 1 constructs rough yard-operation plans, including yard-section choices, allocated quantities, and tentative start and completion times. Stage 2 evaluates these plans through a partially relaxed manpower group configuration and assignment model. If this time-limited model returns a feasible solution $\check{\mathbf{x}}$, its objective value is recorded as the Stage-2 evaluation value $V(\check{\mathbf{x}}) = Obj(\check{\mathbf{x}})$. This value is not the best bound returned by CPLEX and is not treated as a valid lower bound. If $V(\check{\mathbf{x}}) < UB$, Stage 3, Restore, solves the corresponding integer assignment model to recover an integer-feasible solution $\mathbf{x}^I$ and updates the incumbent when a better solution is found. The current sequence is then updated according to relaxed improvement or simulated-annealing acceptance, and the temperature is cooled. After all starts and independent runs finish, LNS-M returns the best restored feasible solution.

4.2. Sequence encoding and initialization

4.2.1. Order-sequence encoding scheme

The LNS solution is represented by a sequence encoding $\chi = \pi$, where $\pi = (\pi_1, \ldots, \pi_{|\mathcal{R}|})$ is a permutation of all orders. This sequence represents the planning priority searched by the outer LNS. It determines the order in which yard sections, operation intervals, and rough driver-demand profiles are constructed in the first stage of the relax-and-restore framework. Because the construction stage updates the remaining yard capacity, and the accumulated rough driver demand after each inserted order, an earlier order in π has more flexibility in selecting yard sections and operation intervals, whereas later orders must adapt to the remaining capacity. The detailed start and completion times are not encoded explicitly. They are selected inside the first stage by the peak-aware scoring rule.

4.2.2. Initial solution generation

For each start, the initial sequence π is generated from the input list. The first start uses the original sequence, and the remaining starts use independently shuffled sequences to diversify the search.

Let N_{repair} denote the limit for additional initialization-repair attempts. The algorithm first applies the first two stages of the relax-and-restore framework to the initial sequence. If Stage 2 finds no feasible solution within its time limit, any CPLEX best bound returned without a feasible solution is not used. The sequence is instead reshuffled, and the first two stages are applied again, up to N_{repair} additional attempts. If all attempts fail, the current start is discarded, and the remaining starts continue. Otherwise, Stage 3, Restore, is applied to the first feasible relaxed solution. If restoration succeeds, the restored integer solution initializes *best* and UB. Otherwise, the start is discarded.

4.3. Three-stage relax-and-restore evaluation framework

Given an encoded order sequence, LNS-M evaluates it through a three-stage relax-and-restore framework. The first stage constructs a fixed set of rough yard-operation plans. The second stage evaluates this plan set with a partially relaxed manpower group configuration and assignment model and returns a feasible solution $\check{\mathbf{x}}$ and its evaluation value $V(\check{\mathbf{x}})$ if such a solution is found within the Stage-2 time limit. The third stage is not called for every candidate; it is triggered only when this evaluation value indicates improvement potential relative to the incumbent integer solution. In this conditional restoration stage, LNS-M fixes the same rough plan set and solves the corresponding integer assignment model with time-indexed group-assignment variables. Its output is an integer-feasible solution $\mathbf{x}^I$ or an infeasibility flag. Thus, the first two stages provide fast construction and heuristic screening in every LNS iteration, whereas Restore provides selective integer-feasibility recovery for promising candidates.

4.3.1. Stage 1: Heuristic rough-plan construction

The first stage, summarized in Algorithm 2, scans orders according to the LNS sequence π. For each order, yard sections are scanned greedily according to operational efficiency or fixed by the known initial yard distribution, and the plan quantity is determined by the remaining order demand and residual yard capacity. The candidate set therefore mainly consists of alternative planned start and completion times for the current order–section pair. The generated rough plan is denoted by

$$p = (r_p, i_p, s_p, e_p, o_p, h_p, Q_p, A_p),$$

where r_p is the order, i_p is the yard section, s_p and e_p are the planned start and completion times, o_p is the first yard-occupation time, and h_p is the last yard-occupation time. For import plans, $o_p = s_p$, and $h_p = b_{r_p}$. For export plans, $o_p = a_{r_p}$, and $h_p = e_p$.

Algorithm 2. Relax-and-restore Stage 1: Heuristic rough-plan construction.

Require: Order sequence π
Ensure: Rough plan set $\mathcal{P}$, or an infeasibility flag

```
Initialize Y_it^rem ← Y_it, d_t ← 0, and P ← ∅
for all orders r following the sequence π do
    Set the remaining order demand according to N_r or ȳ_ri
    while order r still has unassigned quantity do
        C_r ← ∅
        For imports, scan yard sections in descending β_ri and select the first one admitting a feasible
        plan; for fixed initial yard allocations, use the prescribed section
        Set Q_p by the remaining demand and the minimum residual capacity over the candidate
        occupation interval
        Enumerate feasible start and completion times under time-window and capacity restrictions
        for all candidate rough plans p,  do
            Construct the temporal operation distribution A_p
            Evaluate Score(p) by Eq. (4.1)
            Add p to C_r
        end for
        if C_r = ∅ then
            return infeasible
        end if
        Sort C_r by increasing Score(p) and select uniformly from its best min{K_cand, |C_r|} candidates
        Add the selected plan p to P
        Update Y_it^rem, d_t, and the remaining demand of order r
    end while
    if some Y_it^rem < 0 then
        return infeasible
    end if
end for
return P
```

The assigned quantity is denoted by Q_p and represents the number of cars allocated to a specific rough plan p. An order may be split across multiple yard sections, thereby generating several distinct rough plans. In Stage 1, each order greedily occupies the available capacity of a chosen yard section. For orders with pre-existing yard distributions, these quantities are predetermined ($\bar{y}_{ri}$).

The temporal operation distribution is denoted by $A_p = (A_{pt})_{t\in\mathcal{T}}$. The parameter A_{pt} represents the number of cars handled under plan p at time slot t, as derived from the smoothing rule. The rule first assigns at least one car to every active slot of the candidate interval when possible and then allocates the remaining cars one by one to the time slot that causes the smallest projected increase in global driver demand. Ties are broken by the projected driver load, the effective capacity slack, and the existing global operation load. This mechanism encourages order operations to use their time windows without blindly lengthening every order into the same high-overlap periods.

The main design choice is the peak-aware scoring mechanism for alternative start and completion time plans. It follows the stretch–compress trade-off caused by group configuration. For one order–section pair, a longer operation interval is often better because it requires fewer manpower groups in each time slot. However, when several pairs are operated at the same time, their operation intervals overlap, and the peak demand for manpower groups may increase.

Let d_t be the driver demand already assigned at time t. For a candidate plan p, the driver demand added by this plan at time t is defined as follows:

$$n_{pt} = \left\lceil \frac{A_{pt}}{\beta_{r_p i_p}} \right\rceil.$$

Let c_{pt} be the number of plans operated at time t after adding p. Let $\phi(\tilde{d}_{pt})$ be the smallest manpower-group capacity that can cover the driver demand $\tilde{d}_{pt}$ under the group-size limits. For scoring, this demand is defined as the larger of two estimates as follows:

$$\tilde{d}_{pt} = \max\{d_t + n_{pt},\ c_{pt} \min_{k\in\mathcal{K}} S_k\},$$

where $d_t + n_{pt}$ is the driver demand after adding plan p, and $c_{pt} \min_{k\in\mathcal{K}} S_k$ gives each plan operated at time t at least one group of the smallest size. Taking the maximum avoids a too-low estimate when many plans are operated at the same time.

Let $\mathcal{T}_p^{\text{sim}} = \{t \in \mathcal{T} : c_{pt} > 1\}$ be the set of time slots in which more than one plan is operated after adding p. The candidate score is

$$Score(p) = w_1 \max_{t\in\mathcal{T}} \phi(\tilde{d}_{pt}) + w_2 \sum_{t\in\mathcal{T}} \phi(\tilde{d}_{pt})^2 + w_3 \sum_{t\in\mathcal{T}_p^{\text{sim}}} c_{pt}^2, \tag{4.1}$$

where the first term penalizes the largest manpower-group demand over time, the second term smooths the demand profile, and the third term penalizes the number of order–section plans operated at the same time. Whereas Stage 1 prioritizes yard sections with higher operational efficiency, the scoring mechanism balances the resulting manpower demand over time.

4.3.2. Stage 2: Partially relaxed manpower group configuration and assignment evaluation

After the first stage fixes the rough plan set $\mathcal{P}$, the second stage no longer changes yard sections, quantities, or operation intervals. It only determines the manpower group configuration and the time-indexed assignment to the operation plans in this fixed set. Let $\mathcal{P}^O \subseteq \mathcal{P}$ denote the plans that require loading or unloading operations. Static export-occupation plans are excluded from $\mathcal{P}^O$ because they consume yard capacity but require no manpower group assignment in the current shift.

For each operation plan $p \in \mathcal{P}^O$, let $[s_p, e_p]$ be its operation interval, Q_p its operation quantity, and $\beta_p = \beta_{r_p i_p}$ its operational efficiency. The assignment model is formulated as follows:

$$\min \quad \sum_{k \in \mathcal{K}} CTS_k g_k, \tag{4.2}$$

$$\text{s.t.} \quad \sum_{p \in \mathcal{P}^O : s_p \le t \le e_p} z_{ptk} \le g_k, \qquad \forall k \in \mathcal{K},\ t \in \mathcal{T}, \tag{4.3}$$

$$\sum_{t=s_p}^{e_p} \sum_{k \in \mathcal{K}} z_{ptk} S_k \beta_p \ge Q_p, \qquad \forall p \in \mathcal{P}^O, \tag{4.4}$$

$$\sum_{k \in \mathcal{K}} z_{ptk} \ge 1, \qquad \forall p \in \mathcal{P}^O,\ t = s_p, \ldots, e_p, \tag{4.5}$$

$$0 \le g_k \le D_k, \quad g_k \in \mathbb{Z}_+, \qquad \forall k \in \mathcal{K}. \tag{4.6}$$

The objective function (4.2) minimizes the total cost of the configured standardized manpower groups for the fixed rough plan set. Constraint (4.3) ensures that, in each time slot, the number of type-k groups assigned to all active operation plans does not exceed the configured number g_k. Constraint (4.4) requires that the total operation capacity assigned over the operation interval of each operation plan covers its quantity Q_p. Constraint (4.5) enforces continuous operation by requiring at least one manpower group to be assigned to each operation plan in every active time slot. Constraint (4.6) limits the number of configured groups by the available group quantity and keeps the group-configuration decision as an integer.

In the second stage, z_{ptk} is relaxed to a continuous variable as follows:

$$0 \le z_{ptk} \le D_k, \qquad z_{ptk} \in \mathbb{R}_+, \quad \forall p \in \mathcal{P}^O,\ t = s_p, \ldots, e_p,\ k \in \mathcal{K}. \tag{4.7}$$

Constraint (4.7) defines the relaxed domain of the time-indexed assignment variable. It allows the second stage to estimate how the configured manpower groups can be distributed across plans and time slots without enforcing assignment integrality. The model keeps the g_k integer even in the relaxed evaluation version. This is important because g_k represents the actual number of configured manpower groups and appears directly in the objective. If g_k were also relaxed, the evaluation would underestimate the true manpower group cost too aggressively. In contrast, the z_{ptk} variables are much more numerous and are mainly used to distribute the configured manpower groups over plans and time slots. Relaxing z_{ptk} reduces the computational burden of candidate evaluation.

Stage 2 is solved with a time limit of τ_D seconds. If CPLEX identifies a feasible solution $\check{\mathbf{x}}$ within this limit, the algorithm records $V(\check{\mathbf{x}}) = Obj(\check{\mathbf{x}})$ and uses it to evaluate the candidate sequence. If no feasible solution is found, the candidate evaluation is regarded as unsuccessful. During an LNS iteration, such a candidate is rejected, and the search proceeds to the next iteration.

4.3.3. Stage 3: Integer feasibility restoration

The third stage first enforces the integrality constraint on the time-indexed manpower group assignment, and then solves the updated assignment model using CPLEX, as in Stage 2. The integrality constraint is given as follows:

$$z_{ptk} \in \mathbb{Z}_{+}. \tag{4.8}$$

This stage is triggered by a heuristic screening rule. Let $UB = Obj(best)$ denote the objective value of the best integer-feasible solution restored so far. Because the problem is a minimization problem, UB is a valid upper bound. Let $V(\check{\mathbf{x}})$ denote the Stage-2 evaluation value of candidate $\check{\mathbf{x}}$. Restore is called only if

$$V(\check{\mathbf{x}}) < UB. \tag{4.9}$$

If the condition is not satisfied, the algorithm skips the integer restoration step and continues the LNS iterations. Because $V(\check{\mathbf{x}})$ is the objective value of a time-limited Stage-2 feasible solution rather than a valid lower bound, this condition is a heuristic screening rule and does not certify that the skipped candidate cannot improve UB. It is used to limit calls to the more expensive restoration model, which is solved with a time limit τ_R. When restoration succeeds, the candidate obtains an integer-feasible manpower group assignment and a true objective value and can then update the incumbent solution.

4.4. Large-neighborhood operators and acceptance criterion

The LNS-M implementation uses a single random destroy–repair mechanism so that the search behavior is governed by the neighborhood size and the three-stage relax-and-restore evaluation framework.

Random removal. At each iteration, the random-removal step removes q orders from the current order sequence, where

$$q = \max\{1, \min(\lceil \lambda_D |\mathcal{R}| \rceil, |\mathcal{R}| - 1)\},$$

and λ_D is the removal fraction. The positions to be removed are sampled uniformly without replacement from the current sequence. The remaining orders keep their relative order and form a partial sequence. A positive removal fraction allows one LNS move to rebuild a substantial part of the order sequence while preserving part of the incumbent structure.

Random insertion. The random-insertion step shuffles the removed orders and reinserts each of them into a uniformly selected position of the partial sequence. This operation produces a new order sequence π' without relying on problem-specific insertion scores. The randomness helps diversify the search, whereas the first two stages of the relax-and-restore evaluation framework subsequently filter out candidate sequences that cannot generate feasible or promising relaxed plans.

Acceptance criterion. The simulated-annealing rule is applied to the Stage-2 evaluation values of candidates for which feasible solutions are returned. If the candidate improves the current Stage-2 evaluation value, it is accepted directly. Otherwise, it is accepted with probability $P_{\text{acc}} = \exp(-(V(\check{\mathbf{x}}') - V(\check{\mathbf{x}}^c))/Temp)$, where $\check{\mathbf{x}}'$ is the Stage-2 incumbent for the candidate sequence, $\check{\mathbf{x}}^c$ is the Stage-2 incumbent for the current sequence, and $Temp$ is gradually reduced by the cooling factor α until it reaches $Temp_{\min}$. At high temperatures, the algorithm has a higher probability of accepting a worse evaluated candidate and can therefore escape local optima. As the temperature decreases, the search becomes more conservative and gradually concentrates around high-quality regions. The best reported solution is still updated only after Restore succeeds.

5. Computational experiments

In this section, we conduct numerical experiments to validate the effectiveness of the proposed model and LNS-based relax-and-restore matheuristic. The algorithms are implemented in C++, and the experiments are conducted on a personal computer with an 11th Gen Intel(R) Core(TM) i7-1195G7 CPU at 2.90 GHz and 16 GB RAM. The ILP of our problem and the assignment models used in Restore are solved by IBM ILOG CPLEX 12.10. Section 5.1 describes the instance generation and the main algorithmic parameters. Section 5.2 compares LNS-M with CPLEX, a rule-based heuristic (RBH); see Supplementary Section C, and the adaptive LNS-M (ALNS-M) (see Supplementary Section E). Section 5.3 validates the key algorithmic components, including the peak-aware scoring mechanism and the relax-and-restore evaluation framework. Finally, Section 5.4 conducts sensitivity analyses to examine model robustness and derive operational insights.

5.1. Instance generation and parameter setting

5.1.1. Configuration of numerical experiments

The test instances are generated by varying two scale dimensions, namely the number of orders $|\mathcal{R}|$ and the number of yard sections $|\mathcal{I}|$, while keeping a common planning-horizon discretization and a common manpower group-type menu across all instances. Four order scales are considered: $|\mathcal{R}| \in \{10, 20, 50, 100\}$, corresponding to tiny-, small-, medium-, and large-scale instances, respectively. The number of yard sections varies with the order scale. For tiny instances, $|\mathcal{I}| \in \{15, 30\}$. For small and medium instances, $|\mathcal{I}| \in \{30, 50\}$. For large instances, $|\mathcal{I}| \in \{50, 100\}$. The instance configurations are summarized in Table 4.

Table 4. Instance configurations.

Scale	Label	Number of orders $\lvert\mathcal{R}\rvert$	Number of yard sections $\lvert\mathcal{I}\rvert$
Tiny	T	10	15, 30
Small	S	20	30, 50
Medium	M	50	30, 50
Large	L	100	50, 100

Each computational instance is denoted by X-Ib-Na, where X is the scale label in Table 4, Ib is the yard-section-level index, and Na is the random instance index with $a \in \{1, 2, 3\}$. For each scale, I1 and I2 correspond to different yard-section settings listed in Table 4, respectively. For example, T-I1-N2 denotes the second random tiny instance under the setting with 15 yard sections.

Following the operational practices of a major automotive RO-RO terminal in Shanghai, where a 24-hour workday is evenly divided into three shifts, we set the planning horizon for all instances as one 8-hour shift. The horizon is discretized into $|\mathcal{T}| = T = 16$ time slots, and each time slot represents 30 minutes.

Following the parameter settings in Zhang et al. [9], the demand for each order, denoted by N_r, is randomly generated within the range of $[1, 1000]$. The operational efficiency between berth and yard sections is randomly generated from the continuous interval $[0.5, 5.5]$ and subsequently rounded up to the nearest integer. This parameter defines the average number of cars a single driver can transport per time slot on a specific berth–yard route.

Manpower resources are categorized into standardized group types. Following the practical ranges discussed by Mattfeld and Kopfer [1], we assume that each manpower group contains between 5 and 50 drivers. In this study, $\mathcal{K} = 10$ standardized manpower group types are considered, with the set of group sizes defined as $S_k \in \{5, 8, 10, 12, 15, 20, 25, 30, 40, 50\}$. In practice, these distinct group sizes directly correspond to the seating capacities of various minibuses. For each type k, the maximum number of available manpower groups is limited to $D_k = 5$.

5.1.2. Algorithmic parameter setting

The LNS-M algorithm requires several user-controlled parameters. The neighborhood fraction and simulated-annealing parameters are calibrated sequentially on three representative instances, S-I2-N3, M-I1-N2, and L-I1-N1, using three independent runs with common random seeds for each setting. The complete calibration procedure and results are reported in Supplementary Section A. The final parameter settings are listed in Table 5.

Table 5. Main parameters of the LNS-based relax-and-restore matheuristic.

Parameter	Description	Value	Section(s)
N_{run}	Independent runs per start	10	4.1
S	Number of starts	20	4.1
N	Maximum LNS iterations	50	4.1
N_{repair}	Additional initial repair trials	150	4.2.2
q, λ_D	Removal size and removal fraction	$\lceil 0.10R \rceil$	4.4
$Temp_0, \alpha, Temp_{\min}$	Simulated annealing parameters	$100/0.88/10^{-4}$	4.4
τ_D	Stage-2 relaxed evaluation time limit	1 s	4.3
τ_R	Stage-3 restoration time limit	1 s	4.3.3
$w_1/w_2/w_3$	Scoring weights	$1/10^{-4}/0.5$	4.3.1
K_{cand}	Number of top candidates retained for random selection	3	4.3.1

Because the three score components have different intrinsic scales, we use the relative weights $(w_1, w_2, w_3) = (1, 10^{-4}, 0.5)$ to balance their contributions. The first component is a single peak value. The second component sums squared manpower-capacity values over all time slots and therefore has a substantially larger raw magnitude than the other components. In contrast, the third component sums squared counts of concurrent order–section plans, whose numerical values are much smaller than the squared capacity values. To assess local robustness, each weight is independently multiplied by 0.5 and 2, whereas the other two remain fixed. The analysis uses three representative instances and three common-seed runs for each setting. Let $\bar{f}_{hi}$ be the mean objective value for weight setting h on instance i and $\bar{f}_{0i}$ the corresponding baseline value. The reported mean gap is $\Delta_h = (1/3)\sum_{i=1}^{3}((\bar{f}_{hi} - \bar{f}_{0i})/\bar{f}_{0i}) \times 100\%$.

As shown in Table 6, all 63 runs in this analysis return feasible solutions. The mean changes remain within 1.32% of the baseline. Although doubling w_2 yields a marginal average improvement of 0.02%, it ties on S-I2-N3, worsens M-I1-N2 by 0.37%, and improves L-I1-N1 by 0.43%; hence, it does not provide a consistent improvement across instance scales. We therefore retain the baseline relative weights $(1, 10^{-4}, 0.5)$ and conclude that the selected weight ratios are locally robust.

Table 6. Sensitivity analysis of the scoring-weight ratios.

Setting	w_1	w_2	w_3	Mean relative gap Δ_h (%)
Baseline	1.0	1.0×10^{-4}	0.50	0.00
$0.5w_1$	0.5	1.0×10^{-4}	0.50	0.33
$2w_1$	2.0	1.0×10^{-4}	0.50	0.36
$0.5w_2$	1.0	0.5×10^{-4}	0.50	0.44
$2w_2$	1.0	2.0×10^{-4}	0.50	−0.02
$0.5w_3$	1.0	1.0×10^{-4}	0.25	0.88
$2w_3$	1.0	1.0×10^{-4}	1.00	1.32

5.2. *Efficiency assessment for LNS-M*

This subsection compares LNS-M with CPLEX, RBH (Supplementary Section C), and ALNS-M (Supplementary Section E) on 24 test instances. The comparison checks exact-solver consistency, improvement over a fast rule-based constructive benchmark with greedy group configuration and assignment, and whether LNS-M can obtain high-quality solutions without adaptive operator selection. Each LNS-M, RBH, and ALNS-M test is run with 10 random seeds. To provide scale-specific computational budgets, the stopping times of ALNS-M are set with reference to the computational time required by LNS-M: 40, 70, 100, and 150 seconds for tiny, small, medium, and large instances, respectively. These limits are no shorter than the corresponding LNS-M running times.

Table 7. Efficiency assessment of the LNS-based relax-and-restore matheuristic.

Instance	LNS-M				CPLEX			ϕ_C (%)	ϕ_{LB} (%)	RBH		ϕ_R (%)	ALNS-M			ϕ_A^B (%)	ϕ_A^M (%)
	$\check{f}$	$\bar{f}$	ϕ (%)	$\bar{t}$ (s)	f	LB	t (s)			$\bar{f}^R$	$\bar{t}^R$ (s)		$\check{f}^A$	$\bar{f}^A$	$\bar{t}^A$ (s)		
T-I1-N1	4160.00	4160.00	0.00	7.28	4160.00	4160.00	33.85	0.00	0.00	4800.00	2.32	-13.33	4160.00	4160.00	40.00	0.00	0.00
T-I1-N2	6720.00	6720.00	0.00	21.74	6720.00	6720.00	580.96	0.00	0.00	8192.00	2.88	-17.97	6720.00	6720.00	40.00	0.00	0.00
T-I1-N3	7360.00	7360.00	0.00	27.91	7360.00	7360.00	232.47	0.00	0.00	9376.00	2.43	-21.50	7360.00	7360.00	40.00	0.00	0.00
T-I2-N1	5440.00	5472.00	0.59	12.27	5440.00	5440.00	238.93	0.59	0.59	6336.00	2.91	-13.64	5440.00	5440.00	40.00	0.00	-0.58
T-I2-N2	6400.00	6400.00	0.00	15.42	6400.00	6400.00	306.99	0.00	0.00	7936.00	3.99	-19.35	6400.00	6400.00	40.00	0.00	0.00
T-I2-N3	4800.00	4800.00	0.00	9.90	4800.00	4800.00	326.46	0.00	0.00	6272.00	2.02	-23.47	4800.00	4800.00	40.00	0.00	0.00
S-I1-N1	10880.00	10880.00	0.00	62.22	10880.00	10560.00	3600.00	0.00	3.03	13440.00	3.16	-19.05	10880.00	10880.00	70.00	0.00	0.00
S-I1-N2	9600.00	9888.00	3.00	48.68	9600.00	9600.00	2647.65	3.00	3.00	12000.00	5.74	-17.60	9600.00	9760.00	70.00	0.00	-1.29
S-I1-N3	8320.00	8576.00	3.08	34.10	8960.00	7853.33	3600.00	-4.29	9.20	10496.00	5.19	-18.29	8320.00	8352.00	70.00	0.00	-2.61
S-I2-N1	10560.00	10560.00	0.00	40.97	10560.00	10560.00	3094.46	0.00	0.00	12832.00	7.44	-17.71	10560.00	10560.00	70.00	0.00	0.00
S-I2-N2	9600.00	9600.00	0.00	37.30	9600.00	8618.67	3600.00	0.00	11.39	10912.00	4.98	-12.02	9600.00	9600.00	70.00	0.00	0.00
S-I2-N3	10560.00	10560.00	0.00	51.56	11840.00	10048.00	3600.00	-10.81	5.10	12928.00	6.52	-18.32	10560.00	10560.00	70.00	0.00	0.00
M-I1-N1	26560.00	26976.00	1.57	68.36	33920.00	25586.06	3600.00	-20.47	5.43	34496.00	11.18	-21.80	26880.00	27328.00	100.00	1.20	1.30
M-I1-N2	28480.00	28736.00	0.90	59.57	37440.00	26960.67	3600.00	-23.25	6.58	37408.00	10.37	-23.18	28480.00	28864.00	100.00	0.00	0.45
M-I1-N3	33600.00	33984.00	1.14	73.44	50880.00	32067.00	3600.00	-33.21	5.98	44384.00	12.58	-23.43	33920.00	34592.00	100.00	0.95	1.79
M-I2-N1	26560.00	27264.00	2.65	73.73	35200.00	25634.00	3600.00	-22.55	6.36	35840.00	14.93	-23.93	27200.00	27520.00	100.00	2.41	0.94
M-I2-N2	27840.00	28096.00	0.92	61.39	37120.00	26664.67	3600.00	-24.31	5.37	36672.00	13.43	-23.39	27840.00	28320.00	100.00	0.00	0.80
M-I2-N3	25920.00	26112.00	0.74	72.86	33600.00	24624.00	3600.00	-22.29	6.04	34368.00	11.27	-24.02	25920.00	26112.00	100.00	0.00	0.00
L-I1-N1	48320.00	49280.00	1.99	99.44	105920.00	44502.23	3600.00	-53.47	10.74	65600.00	31.64	-24.88	49280.00	50112.00	150.00	1.99	1.69
L-I1-N2	49920.00	51904.00	3.97	67.87	88320.00	47404.00	3600.00	-41.23	9.49	68576.00	23.19	-24.31	51520.00	52992.00	150.00	3.21	2.10
L-I1-N3	57600.00	59264.00	2.89	32.38	–	53522.54	3600.00	–	10.73	92000.00	10.77	-35.58	56000.00	57408.00	150.00	-2.78	-3.13
L-I2-N1	52480.00	54272.00	3.41	65.98	–	49590.67	3600.00	–	9.44	72320.00	28.23	-24.96	54400.00	55200.00	150.00	3.66	1.71
L-I2-N2	46720.00	48384.00	3.56	71.69	–	43896.67	3600.00	–	10.22	65376.00	28.06	-25.99	47680.00	49024.00	150.00	2.05	1.32
L-I2-N3	57600.00	59552.00	3.39	70.96	–	54331.33	3600.00	–	9.61	79040.00	26.47	-24.66	57920.00	60640.00	150.00	0.56	1.83

Table 7 reports the best objective value $\check{f}$, average objective value $\bar{f}$, and average running time $\bar{t}$ of LNS-M, together with its stability gap $\phi = (\bar{f} - \check{f})/\check{f} \times 100\%$. The sample standard deviations and 95% confidence intervals of the ten LNS-M runs are reported in Supplementary Section B. For CPLEX, f, LB, and t denote the feasible objective value, lower bound, and running time, with $\phi_C = (\bar{f} - f)/f \times 100\%$ and $\phi_{LB} = (\bar{f} - LB)/LB \times 100\%$. For RBH, $\bar{f}^R$ and $\bar{t}^R$ denote the average objective value and average running time, with $\phi_R = (\bar{f} - \bar{f}^R)/\bar{f}^R \times 100\%$. For ALNS-M, $\check{f}^A$, $\bar{f}^A$, and $\bar{t}^A$ denote the best objective value, average objective value, and average running time, with $\phi_A^B = (\check{f}^A - \check{f})/\check{f} \times 100\%$ and $\phi_A^M = (\bar{f}^A - \bar{f})/\bar{f} \times 100\%$. A dash (−) indicates that the corresponding value is unavailable within the time limit.

The following observations can be drawn from Table 7.

1) Comparison with CPLEX feasible solutions and lower bounds: For the six tiny instances, CPLEX obtains optimal solutions within the time limit, and the best solutions found by LNS-M match these objective values. Based on the ten-run LNS-M averages, the mean ϕ_C and ϕ_{LB} are both 0.10% for the tiny instances. This validates the correctness and consistency of the LNS-M solution process on small benchmark cases. As the instance size increases, CPLEX becomes much less effective within the 3600-second time limit. It reports feasible solutions for all small and medium instances, but these solutions are often worse than those found by LNS-M: The average ϕ_C is −2.02% for small instances and −24.34% for medium instances. For the large instances, CPLEX reports feasible solutions for only two of the six cases, and their average ϕ_C is −47.35%, whereas LNS-M provides feasible solutions for all six cases. The average lower-bound gap ϕ_{LB} is 5.29% for small instances, 5.96% for medium instances, and 10.04% for large instances. For L-I2-N2 and L-I2-N3, CPLEX finds no feasible solution.

2) Stability and scalability of LNS-M: The internal stability indicator ϕ remains small across the test set, with an average value of 1.41% over all 24 instances. The average running time $\bar{t}$ of LNS-M is 15.75 seconds for tiny instances, 45.81 seconds for small instances, 68.23 seconds for medium instances, and 68.05 seconds for large instances. These results show that the algorithm maintains moderate run-to-run deviations and practical runtimes as the number of orders and yard sections increases.

3) Comparison between LNS-M and RBH: LNS-M outperforms RBH on all 24 instances, with an average $\phi_R = -21.35\%$. The advantage is consistent across all instance scales: The average ϕ_R is −18.21% for tiny instances, −17.16% for small instances, −23.29% for medium instances, and −26.73% for large instances. This comparison shows that the proposed LNS-based matheuristic reduces the average manpower cost by about 21.35% compared with a rule-based heuristic.

4) Comparison between LNS-M and ALNS-M: Even with problem-specific operators and adaptive selection, ALNS-M achieves solution quality close to LNS-M. For average objective values, $\phi_A^M = 0.26\%$ on average, with ALNS-M better on 4 instances, worse on 10, and tied on 10. For best objective values, $\phi_A^B = 0.55\%$ on average, with ALNS-M better on 1 instance, worse on 8, and tied on 15. ALNS-M therefore has a slightly higher average objective value overall, but its solution quality is statistically comparable to that of LNS-M according to a paired two-sided Wilcoxon signed-rank test ($p = 0.0554 > 0.05$).

Overall, Table 7 shows that LNS-M retains exact-solver accuracy on tiny instances and improves over the feasible solutions reported by CPLEX. Compared with RBH, LNS-M spends more computational time but obtains substantially better objective values. The comparison with ALNS-M further shows that LNS-M already obtains high-quality solutions without relying on adaptive operator selection.

5.3. Algorithmic component validation

Further experiments are conducted on four representative instances, S-I1-N3, S-I2-N3, M-I1-N1, and M-I2-N1, to evaluate the main algorithmic components of LNS-M. The first analysis evaluates the effectiveness of the peak-aware scoring mechanism in Stage 1. The second analysis evaluates the relax-and-restore evaluation framework, especially Stage 2, the partially relaxed manpower group configuration and assignment evaluation for candidate screening; and Stage 3, the conditional integer restoration step for feasibility recovery.

5.3.1. Value of the peak-aware scoring mechanism

We conduct a term-by-term ablation of the peak-aware scoring rule in Eq (4.1). The scoring rule contains three weighted components: the largest manpower-group demand term, the demand-profile smoothing term, and the simultaneous-order–section-plan penalty term.

We construct four variants. LNS-M(-w_1) removes the first term and therefore no longer directly penalizes the maximum manpower-group demand over time. LNS-M(-w_2) removes the second term and therefore weakens the smoothing pressure on the demand profile. LNS-M(-w_3) removes the third term and therefore no longer penalizes the number of order–section plans operated at the same time. LNS-M(Random) replaces the complete scoring rule with random start- and completion-time generation.

Let $\bar{f}$ be the average objective value of the complete LNS-M, and let $\bar{f}^{(-)}$ be the average objective value of a component-removed variant. The validation gap is computed as $\phi_S = (\bar{f}^{(-)} - \bar{f})/\bar{f} \times 100\%$. Because the objective is minimized, a positive ϕ_S indicates that removing the scoring component worsens the solution quality. The results are summarized in Table 8.

Table 8. Validation results for the three scoring terms of LNS-M.

Instance	LNS-M		LNS-M(-w_1)			LNS-M(-w_2)			LNS-M(-w_3)			LNS-M(Random)		
	$\bar{f}$	$\bar{t}$	$\bar{f}^{(-)}$	$\bar{t}$	ϕ_S	$\bar{f}^{(-)}$	$\bar{t}$	ϕ_S	$\bar{f}^{(-)}$	$\bar{t}$	ϕ_S	$\bar{f}^{N}$	$\bar{t}$	ϕ_N
S-I1-N3	8416.00	34.07	11488.00	18.90	36.50	8640.00	22.58	2.66	8928.00	17.05	6.08	14688.00	30.84	74.52
S-I2-N3	10560.00	34.27	13248.00	22.58	25.45	10560.00	44.35	0.00	10560.00	28.96	0.00	16544.00	20.62	56.67
M-I1-N1	27232.00	66.15	33600.00	34.99	23.38	27744.00	61.91	1.88	33600.00	32.20	23.38	65280.00	36.31	139.72
M-I2-N1	27456.00	68.30	34304.00	53.42	24.94	28096.00	68.26	2.33	33728.00	41.43	22.84	62624.00	31.66	128.09
Avg.					27.57			1.72			13.08			116.03

The following observations can be drawn from Table 8. (1) The largest manpower-group demand term is the most important component. Removing it increases the average objective value by 27.57%, and all four instances become worse. (2) The simultaneous-order–section-plan penalty is also useful. Removing it increases the average objective value by 13.08%, with the largest deterioration on the two medium instances. (3) The demand-profile smoothing term has the smallest marginal effect. Removing it increases the average objective value by 1.72%, so it mainly provides a secondary refinement. (4) LNS-M(Random) performs the worst. Replacing the scoring rule with random timing generation increases the average objective value by 116.03%, showing that the peak-aware scoring mechanism is essential for yard allocation and start–completion time generation.

5.3.2. Value of the relax-and-restore evaluation framework

We next conduct three ablation experiments to evaluate the effectiveness of the relax-and-restore framework. LNS-M(-S2) removes only Stage 2: Every rough plan generated by Stage 1 is sent directly to the time-limited integer restoration model in Stage 3. If Stage 3 obtains an integer-feasible solution within its time limit, its objective value is used by the acceptance rule; otherwise, the candidate is treated as a failed restoration. LNS-M(-S3) retains the Stage 2 relaxed model but replaces the time-limited CPLEX restoration in Stage 3 with the greedy integer group-configuration and assignment procedure described in Supplementary Section D. Finally, LNS-G(greedy) removes both Stage 2 and Stage 3 and applies this greedy procedure directly after Stage 1.

Each variant is run ten times with common random seeds. To make all three ablation variants computationally comparable with LNS-M, their wall-clock budgets are set to 40 seconds for small instances and 70 seconds for medium instances. Once the budget is reached, no new iteration is launched; however, an active candidate evaluation is allowed to finish normally before termination. Table 9 reports I_c, the mean number of fully completed iterations over the ten runs rounded to the nearest integer, for each instance. The summary row reports only the average validation gaps rather than an average iteration count. All three ablation columns use the validation gap ϕ_S defined above; hence, a positive value indicates deterioration relative to the complete LNS-M.

Table 9. Validation results for the relax-and-restore evaluation framework.

Instance	LNS-M		LNS-M(-S2)			LNS-M(-S3)			LNS-G		
	$\bar{f}$	$\bar{t}$ (s)	$\bar{f}^{(-)}$	I_c	ϕ_S (%)	$\bar{f}^{(-)}$	I_c	ϕ_S (%)	$\bar{f}^{(-)}$	I_c	ϕ_S (%)
S-I1-N3	8416.00	34.07	8960.00	6	6.46	10496.00	233	24.71	10208.00	455	21.29
S-I2-N3	10560.00	34.27	10688.00	7	1.21	12480.00	254	18.18	12064.00	427	14.24
M-I1-N1	27232.00	66.15	27712.00	21	1.76	34560.00	140	26.91	33728.00	306	23.85
M-I2-N1	27456.00	68.30	28224.00	24	2.80	35424.00	126	29.02	33760.00	263	22.96
Avg.					3.06			24.71			20.59

Table 9 shows that all 120 runs of the three ablation variants return feasible solutions. Removing Stage 2 while retaining time-limited ILP-based Stage 3 restoration increases the average objective value by 3.06%. This modest but consistent deterioration indicates that the relaxed model helps identify promising candidates and avoids spending the limited computational budget on unproductive integer restorations. Replacing Stage 3 with greedy integer restoration causes a much larger average deterioration of 24.71%, showing that ILP-based restoration is critical for converting the potential identified by Stage 2 into coordinated integer group assignments. When both stages are replaced in LNS-G, the average deterioration is 20.59%. LNS-G outperforms LNS-M(-S3) because its broader search compensates for the misalignment between Stage 2's relaxed ranking and greedy restoration quality. Both greedy-restoration variants deteriorate much more than LNS-M(-S2), confirming that Stage 3 is the dominant driver of solution quality. Together, these ablations establish the distinct roles of Stage 2 in screening and Stage 3 in high-quality feasibility recovery.

5.4. Sensitivity analysis for robustness validation and operational insights

This subsection reports three complementary analyses. Section 5.4.1 examines demand, yard-capacity, and time-window scenarios. Section 5.4.2 compares our model based on the yard-capacity reservation strategy with the model based on the dynamic yard-occupancy formulation, and Section 5.4.3 examines manpower-group availability and group-size diversity.

5.4.1. Operational scenario analysis

We use T-I1-N3 and T-I2-N2 to examine three operational factors while holding all other inputs fixed. The demand factor η_N scales order quantities and the associated loading or unloading quantities, the yard factor η_Y scales all time-dependent yard capacities, and the time-window factor η_W scales each effective operation-window width around its midpoint. The baseline is $\eta_N = \eta_Y = \eta_W = 1$. Demand and time-window scenarios use factors 0.8 and 1.2, whereas the yard-capacity analysis focuses on the capacity-sensitive range $\eta_Y \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. Each instance–scenario combination is solved by CPLEX with a 3600-second time limit and a target optimality gap of zero. Let f and LB denote the incumbent feasible objective value and the best valid lower bound, respectively. We report the CPLEX optimality gap as $(f - LB)/f \times 100\%$. Let f_i^0 denote the baseline objective value (7360 for T-I1-N3 and 6400 for T-I2-N2); the objective change is $(f - f_i^0)/f_i^0 \times 100\%$. A dash indicates that the corresponding value is unavailable.

As shown in Table 10, CPLEX finds a feasible solution in every demand and time-window scenario and proves optimality in six of the eight cases. It reaches the time limit with gaps of 10.00% for T-I1-N3 under $\eta_N = 0.8$ and 4.35% for T-I2-N2 under $\eta_N = 1.2$. Based on the incumbent objectives, reducing demand to $\eta_N = 0.8$ decreases the objective values by 13.04% and 15.00% for T-I1-N3 and T-I2-N2, respectively, whereas increasing demand to $\eta_N = 1.2$ raises them by 17.39% and 15.00%. Tightening the time windows to $\eta_W = 0.8$ raises the optimal objective values by 13.04% and 15.00%, whereas relaxing them to $\eta_W = 1.2$ leaves T-I1-N3 unchanged and reduces T-I2-N2 by 10.00%. Overall, higher demand and tighter time windows increase manpower requirements, and the reported bounds quantify the remaining uncertainty in the two time-limited solutions.

Table 10. CPLEX results for the demand and time-window scenarios.

Instance	Factor	Level	Status	f	LB	CPLEX gap (%)	Obj. change (%)
T-I1-N3	η_N	0.8	Feasible	6400	5760	10.00	−13.04
T-I1-N3	η_N	1.2	Optimal	8640	8640	0.00	17.39
T-I1-N3	η_W	0.8	Optimal	8320	8320	0.00	13.04
T-I1-N3	η_W	1.2	Optimal	7360	7360	0.00	0.00
T-I2-N2	η_N	0.8	Optimal	5440	5440	0.00	−15.00
T-I2-N2	η_N	1.2	Feasible	7360	7040	4.35	15.00
T-I2-N2	η_W	0.8	Optimal	7360	7360	0.00	15.00
T-I2-N2	η_W	1.2	Optimal	5760	5760	0.00	−10.00

Table 11 reports the yard-capacity results. The CPLEX results reveal a capacity threshold. Both instances attain their baseline objective values from $\eta_Y = 0.3$ onward. For T-I1-N3, contraction to $\eta_Y = 0.2$ and 0.1 raises the incumbent objective by 8.70% and 21.74%, respectively. These runs terminate at the time limit with optimality gaps of 8.00% and 13.41%. For T-I2-N2, CPLEX proves infeasibility at $\eta_Y = 0.1$ and 0.2. Thus, moderate capacity reductions have no cost effect in these instances, whereas severe contraction either increases manpower cost or eliminates feasibility.

Table 11. Yard-capacity scenario results.

Instance	η_Y	Status	f	LB	CPLEX gap (%)	Obj. change (%)
T-I1-N3	0.1	Feasible	8960	7758.19	13.41	21.74
T-I1-N3	0.2	Feasible	8000	7360	8.00	8.70
T-I1-N3	0.3	Optimal	7360	7360	0.00	0.00
T-I1-N3	0.4	Optimal	7360	7360	0.00	0.00
T-I1-N3	0.5	Optimal	7360	7360	0.00	0.00
T-I2-N2	0.1	Infeasible	–	–	–	–
T-I2-N2	0.2	Infeasible	–	–	–	–
T-I2-N2	0.3	Optimal	6400	6400	0.00	0.00
T-I2-N2	0.4	Optimal	6400	6400	0.00	0.00
T-I2-N2	0.5	Optimal	6400	6400	0.00	0.00

5.4.2. Comparison with dynamic yard occupancy

Our reservation model reserves yard space throughout each order–section occupation interval, whereas the dynamic model (Supplementary Section F) updates car accumulation and release each period. We compare the two models on the ten instance-capacity combinations from Table 11 plus the two baseline cases at $\eta_Y = 1.0$. As shown in Table 12, the two models yield the same feasibility status and, whenever feasible, the same objective value and total driver count.

Table 12. Comparison of dynamic occupancy and capacity reservation in the regular and baseline yard-capacity scenarios.

		Dynamic occupancy			Capacity reservation			
Instance	η_Y	Status	f^D	Drivers	Status	f^R	Drivers	Δf (%)
T-I1-N3	0.1	Feasible	8960	28	Feasible	8960	28	0.00
T-I1-N3	0.2	Feasible	8000	25	Feasible	8000	25	0.00
T-I1-N3	0.3	Optimal	7360	23	Optimal	7360	23	0.00
T-I1-N3	0.4	Feasible	7360	23	Optimal	7360	23	0.00
T-I1-N3	0.5	Feasible	7360	23	Optimal	7360	23	0.00
T-I1-N3	1.0	Feasible	7360	23	Optimal	7360	23	0.00
T-I2-N2	0.1	Infeasible	–	–	Infeasible	–	–	–
T-I2-N2	0.2	Infeasible	–	–	Infeasible	–	–	–
T-I2-N2	0.3	Optimal	6400	20	Optimal	6400	20	0.00
T-I2-N2	0.4	Optimal	6400	20	Optimal	6400	20	0.00
T-I2-N2	0.5	Optimal	6400	20	Optimal	6400	20	0.00
T-I2-N2	1.0	Optimal	6400	20	Optimal	6400	20	0.00

Here, f^D and f^R denote the incumbent objectives of the dynamic and reservation formulations, respectively, and $\Delta f = (f^R - f^D)/f^D \times 100\%$. "Feasible" indicates that CPLEX found an incumbent but did not prove optimality within the time limit, whereas "Infeasible" indicates proven infeasibility. In five feasible cases, both are solved to optimality, establishing identical optima. In the other five, at least one model terminates without an optimality proof; the equal objective values only indicate no observed difference within the time limit, not equivalence of the true optima. Group-size combinations may differ when multiple configurations share the same cost. Thus, these regular scenarios reveal no cost penalty from full-interval reservation but do not establish general equivalence. We next construct a capacity-critical instance to expose the mechanism and quantify the potential cost when both models are solved to optimality.

To isolate the mechanism behind the potential conservatism of capacity reservation, we construct a capacity-critical instance from T-I2-N2 with $\eta_Y = 0.3$. The initial distribution is rearranged so that Export Order 6 occupies all 169 spaces of the high-efficiency section c7. Its operation window is changed to $[0, 3]$, allowing it to release 85 spaces in Period 1 and leave 84 cars in c7 in Period 2. Import Order 0 is assigned the window $[1, 3]$ and requires 83 cars to be unloaded in Period 2. Thus, under dynamic occupancy, the 83 newly unloaded cars can share c7 with the remaining 84 cars because $84 + 83 = 167 \leq 169$. Under capacity reservation, the full 169 spaces remain reserved for Order 6 throughout its occupation interval, so the released capacity cannot be used by Order 0. All other orders serve as background operations. The two formulations use identical instance data and cost parameters; their only difference is the treatment of yard occupancy.

The dynamic formulation configures 10 drivers and obtains an optimal objective value of 3200, whereas the reservation formulation configures 15 drivers and obtains 4800. Thus, in this deliberately capacity-critical case, full-interval reservation increases both configured manpower and the objective value by 50%. Figure 4 shows the corresponding order-level manpower profiles, where the manpower assigned to order r in period t is $W_{rt} = \sum_{i,k} S_k z_{ritk}$. The dynamic formulation permits gradual capacity release and more efficient temporal coordination, but the reservation formulation must avoid relying on immediate interorder sharing of nominally released space and consequently requires a larger manpower configuration.

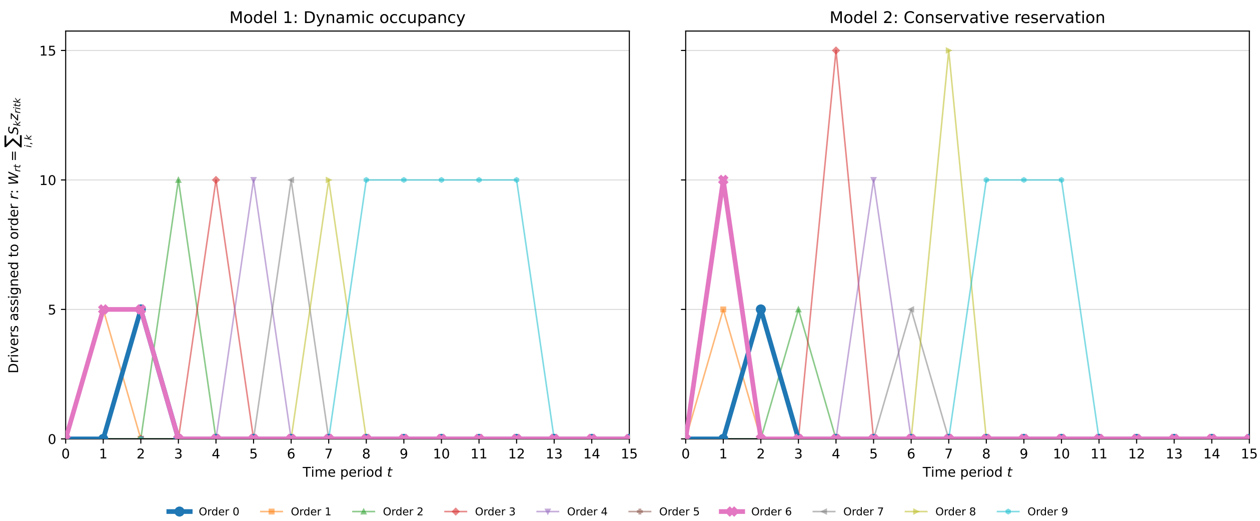


Figure 4. Order-level manpower assignment under dynamic yard occupancy and conservative capacity reservation in the capacity-critical instance.

This comparison clarifies the trade-off underlying the reservation assumption. The dynamic formulation represents physical car flows more precisely and can use yard and manpower resources more efficiently, but its planned gradual accumulation and release trajectories must be reliably implemented through real-time coordination between consecutive operations. The proposed reservation formulation is more conservative and may require additional manpower in capacity-critical settings.

However, it is more suitable for tactical planning when such gradual car-flow trajectories cannot be reliably implemented: Released spaces may be fragmented or not immediately reusable, and reserving the assigned capacity avoids plans that depend on idealized instantaneous yard sharing and perfect operational synchronization.

This evidence leads to a conditional yard-planning strategy. When terminal operators cannot reliably implement the planned gradual accumulation and release of cars, they can adopt the conservative reservation strategy to obtain a more implementable tactical plan and avoid reliance on idealized instantaneous space sharing. This also underscores the importance of timely and reliable order information from RO-RO carriers, as such information enables terminal operators to make more flexible yard-allocation decisions.

5.4.3. Manpower-group availability and group-size diversity

We first examine the effect of manpower group availability and group-size removals on instances T-I1-N3 and T-I2-N2. The available group sizes are 5, 8, 10, 12, 15, 20, 25, 30, 40, and 50 drivers per group, corresponding to group types $k = 0, \ldots, 9$. Two types of tests are conducted. First, we impose a common upper limit $\bar{D}$ on all group types and vary it from 5 to 1. Second, starting from the $\bar{D} = 1$ solution of each instance, we remove the group sizes used in that solution one at a time. Thus, the removed group sizes differ by instance: The 5-, 8-, and 10-driver groups are tested for T-I1-N3, whereas the 5- and 15-driver groups are tested for T-I2-N2. Table 13 reports the resulting configurations obtained by CPLEX. Let f^S denote the feasible objective value under a sensitivity scenario, and let f^0 denote the corresponding baseline value under $\bar{D} = 5$. The relative gap is $\phi_S = (f^S - f^0)/f^0 \times 100\%$.

Table 13 shows that increasing the common availability limit from $\bar{D} = 1$ to $\bar{D} = 5$ does not reduce the objective value in either instance, mainly because the available set of group sizes under $\bar{D} = 1$ is already sufficient to satisfy the total driver demand in these tiny-scale cases. Nevertheless, the resulting group configurations may differ. When several configurations have the same manpower cost, the terminal can choose among them according to practical operational preferences, such as using fewer manpower groups within the yard. However, removing group sizes used in the $\bar{D} = 1$ solution increases the objective value by 8.70% for T-I1-N3 and 15.00% for T-I2-N2 because the remaining group types can no longer combine perfectly to meet the required driver demand and therefore lead to over-configuration. These results suggest that maintaining a diverse set of group sizes, rather than simply increasing the number of identical groups, can preserve configuration flexibility and avoid costly over-configuration.

The group-size sensitivity analysis suggests that terminal operators should maintain a diversified set of standardized manpower-group sizes rather than relying on a single configuration. Such diversity improves the ability to accommodate heterogeneous manpower requirements while avoiding unnecessary capacity caused by oversized groups. Accordingly, group configuration and assignment should be considered jointly so that available groups can be matched more closely with the manpower requirements and time windows of individual operations.

Table 13. Sensitivity results for maximum number of manpower groups of each type.

Instance	Scenario	f^S	ϕ_S (%)	Configured manpower groups	Total drivers	t (s)
T-I1-N3	Common limit $\bar{D} = 5$	7360	0.00	$3 \times 5 + 1 \times 8$	23	139.25
T-I1-N3	Common limit $\bar{D} = 4$	7360	0.00	$1 \times 5 + 1 \times 8 + 1 \times 10$	23	299.49
T-I1-N3	Common limit $\bar{D} = 3$	7360	0.00	$1 \times 5 + 1 \times 8 + 1 \times 10$	23	16.23
T-I1-N3	Common limit $\bar{D} = 2$	7360	0.00	$1 \times 5 + 1 \times 8 + 1 \times 10$	23	195.19
T-I1-N3	Common limit $\bar{D} = 1$	7360	0.00	$1 \times 5 + 1 \times 8 + 1 \times 10$	23	196.32
T-I1-N3	No 5-driver group	8000	8.70	$1 \times 10 + 1 \times 15$	25	910.89
T-I1-N3	No 8-driver group	8000	8.70	$1 \times 5 + 1 \times 20$	25	747.44
T-I1-N3	No 10-driver group	8000	8.70	$1 \times 5 + 1 \times 8 + 1 \times 12$	25	13.95
T-I2-N2	Common limit $\bar{D} = 5$	6400	0.00	$2 \times 5 + 1 \times 10$	20	213.35
T-I2-N2	Common limit $\bar{D} = 4$	6400	0.00	4×5	20	1341.32
T-I2-N2	Common limit $\bar{D} = 3$	6400	0.00	$1 \times 5 + 1 \times 15$	20	983.25
T-I2-N2	Common limit $\bar{D} = 2$	6400	0.00	$2 \times 5 + 1 \times 10$	20	512.06
T-I2-N2	Common limit $\bar{D} = 1$	6400	0.00	$1 \times 5 + 1 \times 15$	20	1212.15
T-I2-N2	No 5-driver group	7360 (LB: 7040.00)	15.00	$1 \times 8 + 1 \times 15$	23	3600.36
T-I2-N2	No 15-driver group	7360 (LB: 6914.17)	15.00	$1 \times 5 + 1 \times 8 + 1 \times 10$	23	3600.64

6. Conclusion

This study addresses a joint yard allocation and manpower group configuration and assignment problem in automotive RO-RO terminals. In this problem, transfer drivers are dispatched as standardized manpower groups and are transported by a minibus between vessels and yard sections. We first formulate an integer linear programming model. The model jointly decides yard-section allocation, the planned start and completion times of order–section pairs, manpower group configuration, and time-indexed group assignment. In this way, it links yard occupation, operation duration, and manpower use in one model.

To solve large-scale instances effectively, we develop an LNS-based relax-and-restore matheuristic, named LNS-M. LNS-M uses an order-sequence encoding and a three-stage relax-and-restore evaluation framework. The first stage builds rough yard-operation plans and uses a peak-aware scoring mechanism to decide the start and completion times of each order–section pair. The second stage evaluates these plans with a partially relaxed manpower group configuration and assignment model. The third stage conditionally restores integer-feasible group assignments for promising candidates. Extensive numerical experiments show that the proposed method is effective. LNS-M matches CPLEX optimal solutions on all tiny instances and improves over CPLEX feasible solutions on many small instances and on all medium and large instances. Compared with the rule-based heuristic, LNS-M reduces the average manpower cost by 21.35%. The component analysis also shows that the peak-aware scoring mechanism and the relax-and-restore evaluation framework are important for solution quality. Finally, the sensitivity analysis suggests that maintaining a diverse set of group sizes, rather than simply increasing the number of identical groups, can preserve configuration flexibility and avoid costly over-configuration.

Several meaningful extensions can be considered in future research. First, manpower group configuration can be planned across multiple shifts so that group sizes and group numbers are coordinated over a longer planning horizon. Second, future work can consider uncertainty in order arrivals, processing times, and yard availability, which may change the planned start and completion times of order–section pairs. Finally, a decision support system can be developed to jointly support berth planning, yard allocation, manpower group configuration, and manpower group assignment in RO-RO terminals.

Author contributions

Ziqiao Mei: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing–original draft, visualization. Feng Chen: Conceptualization, resources, formal analysis, writing–review & editing, supervision, project administration. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant number 72172091].

Data availability

The input data can be found in the GitHub: https://github.com/MzqCode/Integrated-yard-and-manpower-group-planning.git

Use of Generative-AI tools declaration

Statement: During the preparation of this work, the authors used GPT-5.6 Sol (OpenAI, 2026) to improve the language and readability of the text written by the authors, and to assist in creating Python scripts for data visualizations. All AI-generated outputs were carefully reviewed, edited, and adapted as needed to ensure accuracy, relevance, and reproducibility. The authors take full responsibility for the content of the published article.

Conflict of interest

The authors declare no conflict of interest.

References

1. D. Mattfeld, H. Kopfer, Terminal operations management in vehicle trans-shipment, *Transp. Res. Part A Policy Pract.,* **37** (2003), 435–452. https://doi.org/10.1016/S0965-8564(02)00035-6
2. J. F. Cordeau, G. Laporte, L. Moccia, G. Sorrentino, Optimizing yard assignment in an automotive trans-shipment terminal, *Eur. J. Oper. Res.,* **215** (2011), 149–160. https://doi.org/10.1016/j.ejor.2011.06.008
3. R. Iannone, S. Miranda, L. Prisco, S. Riemma, Cost management in car terminals, *Proc. Int. Conf. Harbor Marit. Multimodal Logist.,* 2014, 133–140.

4. X. Chen, F. Li, B. Jia, J. Wu, Z. Gao, R. Liu, Optimizing storage location assignment in an automotive Ro-Ro terminal, *Transp. Res. Part B Methodol.,* **143** (2021), 249–281. https://doi.org/10.1016/j.trb.2020.10.009
5. Z. Zhang, H. Fan, Yard space allocation for Roll-on/Roll-off port based on resilient yard template, *Adv. Eng. Inform.,* **69** (2026), 103804. https://doi.org/10.1016/j.aei.2025.103804
6. Z. Mei, F. Chen, D. Zhang, Optimizing manpower-dependent berth allocation problem in automotive terminals, *Transp. Res. Part E Logist. Transp. Rev.,* **211** (2026), 104851. https://doi.org/10.1016/j.tre.2026.104851
7. F. Jaehn, T. Wensing, F. Wiedra, Multi-mode multi-skill personnel scheduling in Roll-on/Roll-off terminals, *Marit. Transp. Res.,* **11** (2026), 100156. https://doi.org/10.1016/j.martra.2026.100156
8. X. Chen, Q. Tian, F. Li, B. Jia, Optimization for the management performance of drivers at automotive Ro–Ro terminals: Consider efficiency and fairness, *Transp. Res. Part E Logist. Transp. Rev.,* **201** (2025), 104221. https://doi.org/10.1016/j.tre.2025.104221
9. D. Zhang, F. Chen, Z. Mei, Optimization on joint scheduling of yard allocation and transfer manpower assignment for automobile RO-RO terminal, *Transp. Res. Part E Logist. Transp. Rev.,* **177** (2023), 103256. https://doi.org/10.1016/j.tre.2023.103256
10. D. Li, H. Jin, Y. Zhang, Dynamic worker allocation in seru production systems with actor–critic and pointer networks, *Eur. J. Oper. Res.,* **324** (2025), 62–74. https://doi.org/10.1016/j.ejor.2025.01.012
11. S. Tian, J. Fan, C. Zhang, X. Li, L. Gao, A matheuristic for assembly operation scheduling problem with flexible workforce configuration, *Comput. Ind. Eng.,* **211** (2026), 111590. https://doi.org/10.1016/j.cie.2025.111590
12. M. Wen, E. Thorlacius, J. Larsen, Do costs and team-building conflict in airline crew scheduling? An individual crew pairing approach, *Transp. Res. Part E Logist. Transp. Rev.,* **197** (2025), 104057. https://doi.org/10.1016/j.tre.2025.104057
13. M. P. Mandal, A. Santini, C. Archetti, Tactical workforce sizing and scheduling decisions for last-mile delivery, *Eur. J. Oper. Res.,* **323** (2025), 153–169. https://doi.org/10.1016/j.ejor.2024.12.006
14. H. Calik, A. Oulamara, C. Prodhon, S. Salhi, The electric location-routing problem with heterogeneous fleet: Formulation and Benders decomposition approach, *Comput. Oper. Res.,* **131** (2021), 105251. https://doi.org/10.1016/j.cor.2021.105251
15. A. Imai, H. C. Chen, E. Nishimura, S. Papadimitriou, The simultaneous berth and quay crane allocation problem, *Transp. Res. Part E Logist. Transp. Rev.,* **44** (2008), 900–920. https://doi.org/10.1016/j.tre.2007.03.003
16. F. Meisel, C. Bierwirth, A framework for integrated berth allocation and crane operations planning in seaport container terminals, *Transp. Sci.,* **47** (2013), 131–147. https://doi.org/10.1287/trsc.1120.0419
17. J. He, Y. Huang, W. Yan, S. Wang, Integrated internal truck, yard crane and quay crane scheduling in a container terminal considering energy consumption, *Expert Syst. Appl.,* **42** (2015), 2464–2487. https://doi.org/10.1016/j.eswa.2014.11.016

18. L. Zhen, S. Yu, S. Wang, Z. Sun, Scheduling quay cranes and yard trucks for unloading operations in container ports, *Ann. Oper. Res.,* **273** (2019), 455–478. https://doi.org/10.1007/s10479-016-2335-9
19. R. T. Cahyono, S. P. Kenaka, B. Jayawardhana, Simultaneous allocation and scheduling of quay cranes, yard cranes, and trucks in dynamical integrated container terminal operations, *IEEE Trans. Intell. Transp. Syst.,* **23** (2022), 8564–8578. https://doi.org/10.1109/TITS.2021.3083598
20. H. Li, J. Gao, L. Zhen, X. He, Berth and yard scheduling optimization for a port with diagonal yard layout, *Flex. Serv. Manuf. J.,* **37** (2025), 473–512. https://doi.org/10.1007/s10696-024-09546-w
21. Y. Wang, R. Zheng, Y. Liu, Joint scheduling of value-added services and automobile distribution in RO–RO terminals: A column generation approach, *Comput. Ind. Eng.,* **212** (2026), 111657. https://doi.org/10.1016/j.cie.2025.111657
22. Y. D. Ko, The airfare pricing and seat allocation problem in full-service carriers and subsidiary low-cost carriers, *J. Air Transp. Manag.,* **75** (2019), 92–102. https://doi.org/10.1016/j.jairtraman.2018.12.003
23. C.-H. Tang, P.-Y. Chu, Airline unit load device dispatch and scale in short-term operations involving skypooling—A new sharing economy in airline industry, *Transp. Res. Part C Emerg. Technol.,* **128** (2021), 103164. https://doi.org/10.1016/j.trc.2021.103164
24. C.-H. Tang, Y.-W. Wang, Transportation outsourcing problems considering feasible probabilities under stochastic demands, *Comput. Oper. Res.,* **126** (2021), 105109. https://doi.org/10.1016/j.cor.2020.105109
25. D. Melder, S. Wang, M. H. Mansoor, J. H. Drake, E. K. Burke, The network level airport slot allocation problem: A new formulation and matheuristic solution approach, *Transp. Res. Part B Methodol.,* **207** (2026), 103428. https://doi.org/10.1016/j.trb.2026.103428
26. R. Gaborit, E. V. D. Hurk, O. A. Nielsen, Y. Jiang, Optimisation of bus timetables: An adaptive large neighbourhood search-based matheuristic with a novel operator weight, *Eur. J. Oper. Res.,* **333** (2026), 38–52. https://doi.org/10.1016/j.ejor.2026.01.032
27. A. Algendi, S. Urrutia, L. M. Hvattum, A matheuristic for the home healthcare routing and scheduling problem with ferry-dependent travel times, *Comput. Oper. Res.,* **188** (2026), 107372. https://doi.org/10.1016/j.cor.2025.107372