

Box Allocation Optimization in Meal Kit Delivery

Thi Minh Thu Nguyen*

Loïc Genest†

Alain Zemkoho‡

September 2025

Abstract

This study introduces the Box Allocation Problem (BAP), a novel optimization challenge in the \$1.4 billion UK meal kit delivery market. BAP involves assigning orders across multiple production facilities to minimize daily recipe variations while adhering to capacity and eligibility constraints over a 15-day planning horizon. We formulate BAP as a mixed-integer linear programming (MILP) problem and systematically compare the performance of the COIN-OR Branch and Cut (CBC) solver with heuristic methods, including Tabu Search and Iterative Targeted Pairwise Swap. Scalability experiment on instances with up to 100,000 orders show that CBC consistently achieves optimal solutions in under two minutes, maintaining optimality even under dynamic conditions with fluctuating factory capacities and changing customer orders. By reducing day-to-day recipe discrepancies, this approach supports more accurate ingredient forecasting, decreases food waste, and improves operational efficiency across multi-factory network. These results provide the first comprehensive solution framework for temporal allocation problems in meal kit delivery operations.

Keywords: recipe box allocation, meal kit delivery, CBC solver, heuristics, production planning, sustainability, supply chain management

Abbreviations: B&B, Branch and Bound; BAP, Box Allocation Problem; BPP, Bin Packing Problem; CBC, COIN-OR Branch and Cut; CVRP, Capacitated Vehicle Routing Problem; ITPS, Iterative Targeted Pairwise Swap; LD, Lead Day; MILP, Mixed-Integer Linear Programming; TS, Tabu Search; WMAPE, Weighted Mean Absolute Percentage Error.

1 Introduction

The UK meal kit delivery market has experienced rapid expansion, driven by consumer demand for convenient, nutritious, and personalized meal solutions. Reaching \$1.4 billion in revenue by 2024, this sector presents a highly competitive environment where both established food companies and technology-driven startups compete for market share [24]. Operational excellence in this segment depends heavily on sophisticated supply chain networks capable of coordinating fresh ingredient procurement, inventory management, and time-sensitive distribution across multiple production facilities. The inherent complexity stems from managing perishable goods with short shelf lives, unpredictable demand patterns, and stringent quality requirements – challenges that transform supply chain optimization from a cost consideration into a strategic imperative for survival [13, 4]. Companies that fail to achieve operational agility in response to demand volatility and supply disruptions face immediate risks to customer satisfaction, food waste costs, and ultimately, business viability [23].

Within this operational landscape emerges the Box Allocation Problem (BAP), a critical optimization challenge that involves strategically distributing recipe boxes (Figure 1) across production facilities to minimize day-to-day variations in recipe allocations. BAP seeks to minimize the Weighted Mean Absolute Percentage Error (WMAPE) of daily recipe distributions while satisfying factory capacity constraint and recipe eligibility restriction [16]. This objective delivers dual operational benefits that directly impact both sustainability and profitability. First, consistent recipe allocations across facilities enable accurate ingredient demand forecasting, reducing over-procurement and food spoilage – a critical advantage given that food waste accounts for approximately 10% of global greenhouse gas emissions [26]. Second, stabilized allocation patterns ensure reliable



Figure 1: Recipe box [27]

*Department of Decision Analytics and Risk, Southampton Business School, University of Southampton, Southampton, United Kingdom. Corresponding author. Email: tmtn1n23@soton.ac.uk

†Data Science Manager, London, United Kingdom.

‡School of Mathematical Sciences, University of Southampton, Southampton, United Kingdom.

ingredient availability across all production sites, minimizing stockouts that can disrupt customer orders and damage brand reputation in this highly competitive market. Despite BAP’s fundamental importance to operational efficiency and environmental sustainability in meal kit delivery, it has received minimal attention in the supply chain optimization literature, representing a significant research gap in network-based production planning.

This study addresses the identified research gap by developing a comprehensive mixed-integer linear programming (MILP) formulation of BAP that captures the interplay between facility capacity constraints and recipe eligibility requirements. We systematically benchmark the COIN-OR Branch and Cut (CBC) solver – an exact optimization method – against two established heuristic approaches: Iterative Targeted Pairwise Swap (ITPS) and Tabu Search (TS). These heuristics were selected based on their proven effectiveness in structurally similar combinatorial optimization problems, particularly the Bin Packing Problem (BPP) and Capacitated Vehicle Routing Problem (CVRP), which share BAP’s fundamental characteristics of resource allocation under capacity constraints [14, 15, 8]. TS demonstrates particular strength in escaping local optima through its adaptive memory structure and strategic exploration of large solution spaces [19, 17], while CBC guarantees global optimality through systematic branch-and-bound enumeration combined with sophisticated cut generation and primal heuristics [18, 3]. Our experimental design encompasses diverse operational scenarios, ranging from 10,000 to 100,000 orders across static and dynamic conditions, including fluctuating factory capacities and evolving customer demand patterns. This comprehensive evaluation framework enables rigorous assessment of each method’s computational efficiency, solution quality, and adaptability to real-world complexities.

Our findings highlight the superior performance of the CBC solver in attaining optimal solutions, even under challenging dynamic conditions. By integrating both theoretical foundations and practical applications of BAP, this study makes a substantial contribution to the supply chain optimization literature while offering actionable insights for meal kit delivery enterprises. It bridges a critical knowledge gap by proposing scalable solutions that enhance operational efficiency and promote environmental sustainability in this rapidly growing segment.

The remainder of this paper is structured as follows: Section 2 introduces the characteristics of BAP and presents its mathematical formulation, situating it in relation to two classic optimization problems. Section 3 outlines the optimization methods employed, covering both exact and heuristic approaches. Section 4 reports the results of numerical experiments, comparing the performance of all methods. Finally, Section 5 discusses the implications of the findings, acknowledges limitations, and identifies promising avenues for future research.

2 Problem Statement

2.1 Problem background

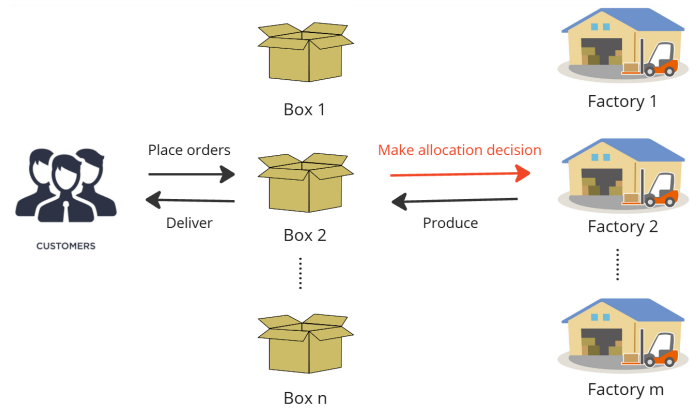
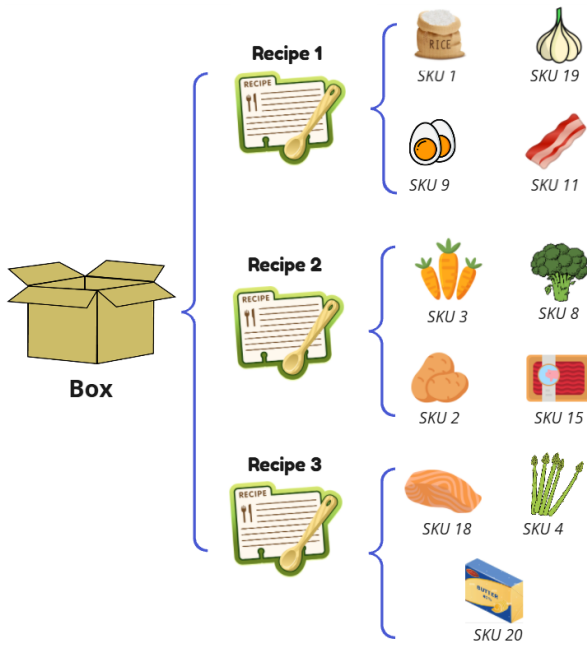
BAP in meal kit delivery involves assigning recipe boxes, structured as in Figure 2a, to multiple factories (Figure 2b). The objective is to minimize daily variations in recipe allocations. For example, if Factory 1 (F1) produced 100 units of Recipe 1 yesterday, the optimal solution would allocate approximately 100 units today.

The problem includes two key constraints. First, each factory (except the final simulated factory m) has a fixed daily production capacity, limiting the number of boxes it can process. To optimize resource use, solutions must fully meet this capacity (Figure 3a). Second, not all factories can produce every recipe, so allocation must follow the eligibility constraint in Figure 3b, ensuring recipes are assigned only to capable facilities.

A key feature of BAP is its temporal dimension, requiring daily allocation of both actual and simulated orders (Figure 4). This study uses a 15-day planning horizon from lead day 18 (LD18) to lead day 3 (LD3). The process begins with total demand forecasting, followed by continuous monitoring of actual orders. When real orders fall short of projections, simulated orders are generated to represent expected demand. Initially, both real and simulated orders are distributed across factories (*soft allocation*). As the planning horizon progresses, actual orders replace simulated ones. By LD3, the allocation consists solely of real orders, transitioning to *hard allocation*, which can no longer be changed. The company then has three days to produce and deliver the boxes. This approach ensures stable daily allocations while adapting to fluctuating demand, enabling a smooth shift from forecasts to actual order fulfillment.

2.2 Mathematical formulation

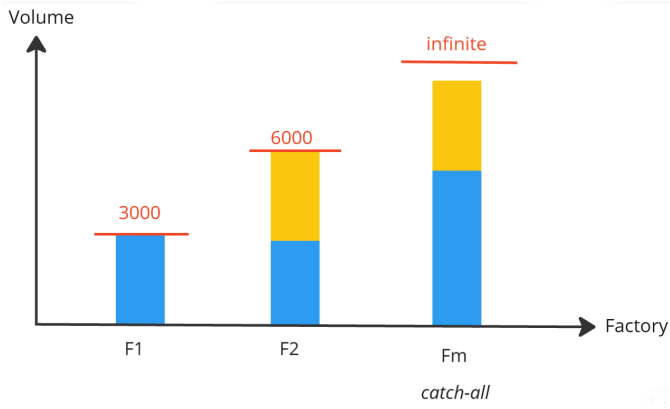
WMAPE is a key forecast accuracy metric used in supply chain management and production planning. It measures the sum of absolute differences between actual and forecasted values, divided by the sum of actual



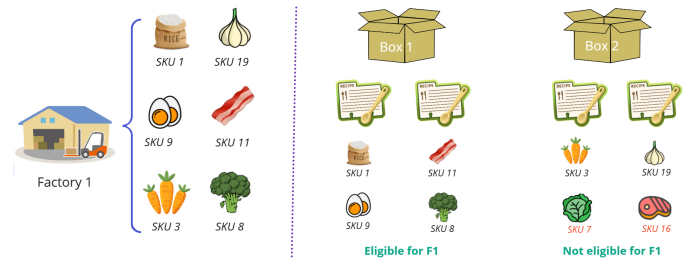
(a) Box structure [16]

(b) Allocation process [16]

Figure 2: Recipe box allocation



(a) Capacity constraint [16]



(b) Eligibility constraint [16]

Figure 3: Main constraints

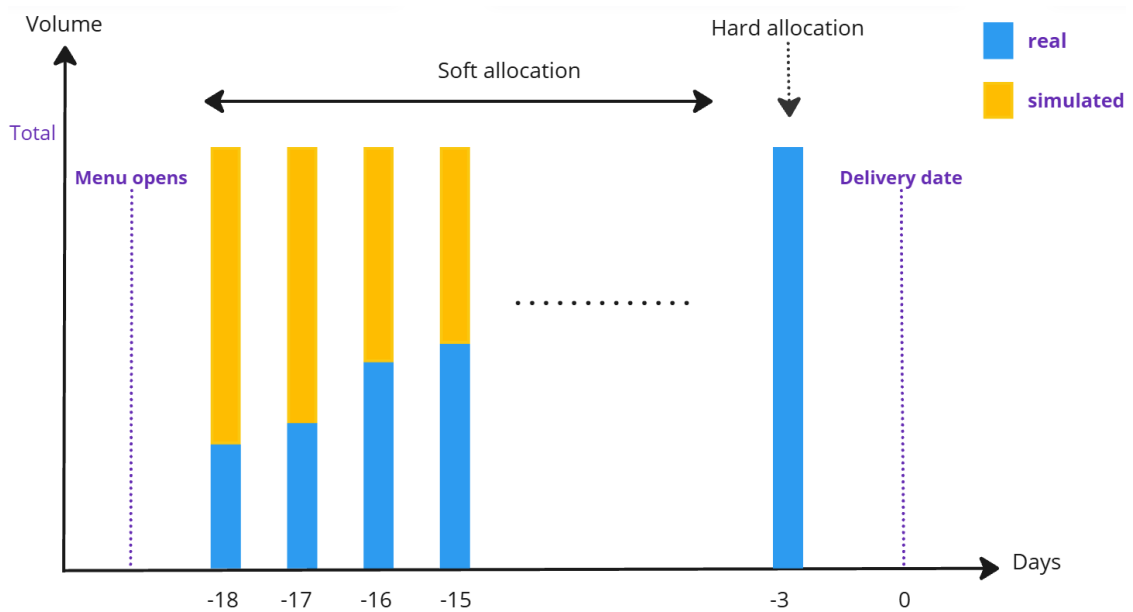


Figure 4: Temporal aspect of BAP [16]

values [7], with lower values indicating higher accuracy. This research focuses on minimizing WMAPE site in Table 1 to provide granular insights and assess production consistency at the factory level.

Table 1: Comparison between WMAPE site and WMAPE global

Aspect	WMAPE site	WMAPE global
Scope	Considers each factory separately	Aggregates data across all factories
Sensitivity	Sensitive to changes in allocation between factories	Not sensitive to changes in allocation between factories
Usefulness	Analyzes stability of production plans at each factory	Provides a broader view of recipe distribution changes

The mathematical formulation of BAP employs the following notation:

- $T = \{-18, -17, \dots, -3\}$: Set of allocation days, indexed by t .
- $I = \{1, 2, \dots, n\}$: Set of recipes, indexed by i .
- $J = \{1, 2, \dots, m\}$: Set of factories, indexed by j .
- $O = \{1, 2, \dots, k\}$: Set of orders, indexed by o .
- $a_{i,t-1}$: Quantity of recipe i at day $t - 1$.
- $a_{i,t}$: Quantity of recipe i at day t .
- $a_{i,j,t}$: Quantity of recipe i allocated to factory j at day t .
- $a_{i,j,t-1}$: Quantity of recipe i allocated to factory j at day $t - 1$.
- $x_{o,j,t}$: Binary decision variable indicating the allocation of order o to factory j on day t .
- $C_{j,t}$: Capacity of factory j on day t .
- $E_{i,j,t}$: Binary parameter indicating the eligibility of recipe i at factory j on day t .

We formulate BAP as a MILP problem:

$$\underset{x}{\text{Minimize}} \quad f(x) = \frac{\sum_{i=1}^n \sum_{j=1}^m |a_{i,j,t} - a_{i,j,t-1}|}{\sum_{i=1}^n a_{i,j,t}} \quad (1)$$

$$\text{subject to} \quad \sum_{o=1}^k x_{o,j,t} = C_{j,t}, \quad \forall j \in J \setminus \{m\}, \quad \forall t \in T, \quad (2)$$

$$x_{o,j,t} \leq \min_{i \in o} E_{i,j,t}, \quad \forall o \in O, \quad \forall j \in J, \quad \forall t \in T, \quad (3)$$

$$\sum_{j=1}^m x_{o,j,t} = 1, \quad \forall o \in O, \quad \forall t \in T, \quad (4)$$

$$x_{o,j,t} \in \{0, 1\}, \quad \forall o \in O, \quad \forall j \in J, \quad \forall t \in T. \quad (5)$$

The objective function (1) minimizes WMAPE site by reducing daily fluctuations in recipe allocations across factories, while accounting for the total recipe volume assigned each day. Recipe quantities $a_{i,j,t}$ are indirectly controlled by the binary decision variable $x_{o,j,t}$ through constraints. When $x_{o,j,t}$ determines which factory j receives order o at time t , it simultaneously establishes how recipes must be distributed. The allocation of an order to a specific factory creates a cascade effect that dictates the required recipe quantities, even though $a_{i,j,t}$ does not explicitly contain $x_{o,j,t}$ in its definition.

Constraint (2) ensures that each factory operates at full capacity to avoid resource waste. Specifically, every factory – except the simulated catch-all factory F_m – must receive order quantities that exactly match its production capacity. Constraint (3) enforces the eligibility requirement, allowing an order to be assigned only to a factory capable of producing all recipes required by that order, thereby preventing infeasible manufacturing assignments. Constraint (4) guarantees that each order is assigned to exactly one factory. Finally, Constraint

(5) requires the allocation variables to be binary (0 or 1), ensuring that each order is either fully assigned to a particular factory or not assigned at all, reflecting the discrete nature of the assignment decisions.

WMAPE global is computed using the following equation:

$$WMAPE_{\text{global}} = \frac{\sum_{i=1}^n |a_{i,t} - a_{i,t-1}|}{\sum_{i=1}^n a_{i,t}}, \quad (6)$$

which establishes a theoretical lower bound to evaluate the solution optimality. When decomposed at the site level, this formula becomes:

$$WMAPE_{\text{global}} = \frac{\sum_{i=1}^n \left| \sum_{j=1}^m a_{i,j,t} - \sum_{j=1}^m a_{i,j,t-1} \right|}{\sum_{i=1}^n a_{i,t}}, \quad (7)$$

By applying the triangle inequality, we derive the fundamental relationship:

$$WMAPE_{\text{global}} \leq \frac{\sum_{i=1}^n \sum_{j=1}^m |a_{i,j,t} - a_{i,j,t-1}|}{\sum_{i=1}^n a_{i,t}} (WMAPE_{\text{site}}) \quad (8)$$

Table 2 presents exemplary order data for two consecutive days, while Table 3 calculates the corresponding WMAPE site and global. This result reveals that despite using the same order data, WMAPE global (0.5) is lower than WMAPE site (1.14), confirming the theoretical relationship in equation (8).

Table 2: Exemplary orders of LD15 and LD14

LD15		
Order ID	Recipe IDs	Assigned Factory
1	1, 10	F1
2	2, 3, 5	F1
3	4, 6	F2
4	2, 3, 7	F2
5	3, 5, 9	F3

LD14		
Order ID	Recipe IDs	Assigned Factory
1	2, 5	F1
2	2, 6, 7	F3
3	3, 5, 9	F1
4	4, 6, 8	F2
5	5, 9, 10	F2

Table 3: WMAPE site and WMAPE global

Recipe	Factory	LD15 ($a_{i,j,t-1}$)	LD14 ($a_{i,j,t}$)	Abs Diff
1	F1	1	0	1
2	F1	1	1	0
2	F2	1	0	1
2	F3	0	1	1
3	F1	1	1	0
3	F2	1	0	1
3	F3	1	0	1
4	F2	1	1	0
5	F1	1	2	1
5	F2	0	1	1
6	F2	1	1	0
6	F3	0	1	1
7	F2	1	0	1
7	F3	0	1	1
8	F2	0	1	1
9	F1	0	1	1
9	F2	0	1	1
9	F3	1	0	1
10	F1	1	0	1
10	F2	0	1	1
SUM			14	16
WMAPE site = 16/14 = 1.14				

Recipe	LD15 ($a_{i,t-1}$)	LD14 ($a_{i,t}$)	Abs Diff
1	1	0	1
2	2	2	0
3	3	1	2
4	1	1	0
5	2	3	1
6	1	2	1
7	1	1	0
8	0	1	1
9	1	2	1
10	1	1	0
SUM		14	7
WMAPE global = 7/14 = 0.5			

The company's ultimate goal is to minimize the area under the curve in Figure 5, which represents the WMAPE site when comparing the final day's allocation with each of the previous days. However, directly optimizing this metric is infeasible due to forecast uncertainties before final day LD3. These uncertainties arise

from uncontrollable factors, including the variability of simulated orders, customer modifications to existing orders, and the continuous arrival of new actual orders throughout the planning period. To tackle this challenge, the company adopts a more practical approach by minimizing allocation variations between consecutive days, as defined in Equation 1. This strategy reframes BAP as a *proxy optimization problem*, where solving a simpler, related problem serves as a substitute for optimizing the more complex objective [28]. By reducing day-to-day allocation discrepancies, this approach effectively lowers cumulative error over the entire planning period, making it a feasible way to achieve the company’s broader strategic goal.

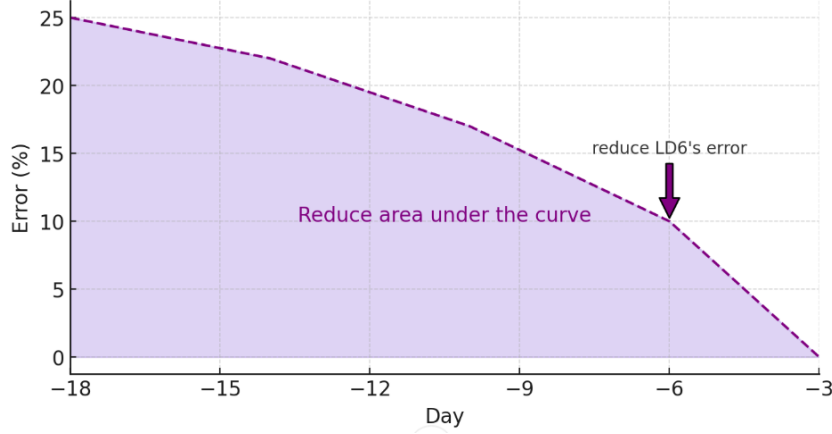


Figure 5: Objective of proxy optimization

BAP shares similarities with two well-known combinatorial optimization problems: BPP and CVRP. These problems rely on a network structure, where nodes represent key entities (factories, storage bins, or vehicles) and edges define decision variables (order allocations, item assignments, or routing sequences). Allocations are subject to capacity constraints to ensure efficient resource utilization, as illustrated in Figure 6. Despite these shared features, each problem has a distinct mathematical formulation. BPP minimizes the number of bins while keeping total item weight or volume within capacity limits. CVRP focuses on reducing total travel costs or distances while ensuring vehicle capacity is not exceeded. BAP introduces the unique objective of minimising temporal variations in recipe allocations across factories.

Their similarities highlight opportunities for cross-domain optimization techniques. The B&B method from BPP has proven effective across various allocation problems [18]. Likewise, heuristic approaches such as 2-Opt and Tabu Search from CVRP have influenced allocation strategies in multiple domains [15, 10]. However, applying these methods to BAP requires substantial modifications to account for its unique constraints and the emphasis on maintaining stable daily allocations. These distinctions place BAP beyond traditional BPP and CVRP models, requiring problem-specific mathematical formulations while strategically leveraging established optimization techniques.

3 Problem Solving

The company seeks an efficient approach to minimize recipe discrepancies between two allocation decisions. Figure 7 illustrates the different swap types employed in heuristics, where blue boxes represent real orders and orange boxes denote simulated orders. To simplify the process, this study adopts the 1:1 swap method. The solving procedure consists of the following steps:

- **Exact method:** CBC solver is directly applied to current orders to determine the optimal allocation that minimizes recipe discrepancies relative to the previous day’s allocation.
- **Heuristics:** The process starts with a greedy algorithm to construct an initial feasible solution. Orders are first allocated to F1, prioritizing those with fewer eligible factories (F1-F3 first, then F1-F2-F3) until capacity is reached. Remaining orders with eligible factories F1-F2-F3 and F2-F3 are allocated to F2 until its capacity is full. Any unallocated orders are assigned to F3, the catch-all factory without constraints. ITPS and TS are then applied to refine the initial allocation.
- **Evaluation:** The performance and scalability of each method are assessed by comparing the final WMAPE site against the WMAPE global and analyzing optimization times for varying order quantities. The most effective method is subsequently applied in temporal tests to evaluate performance over a 15-day planning period and its adaptability to dynamic changes.

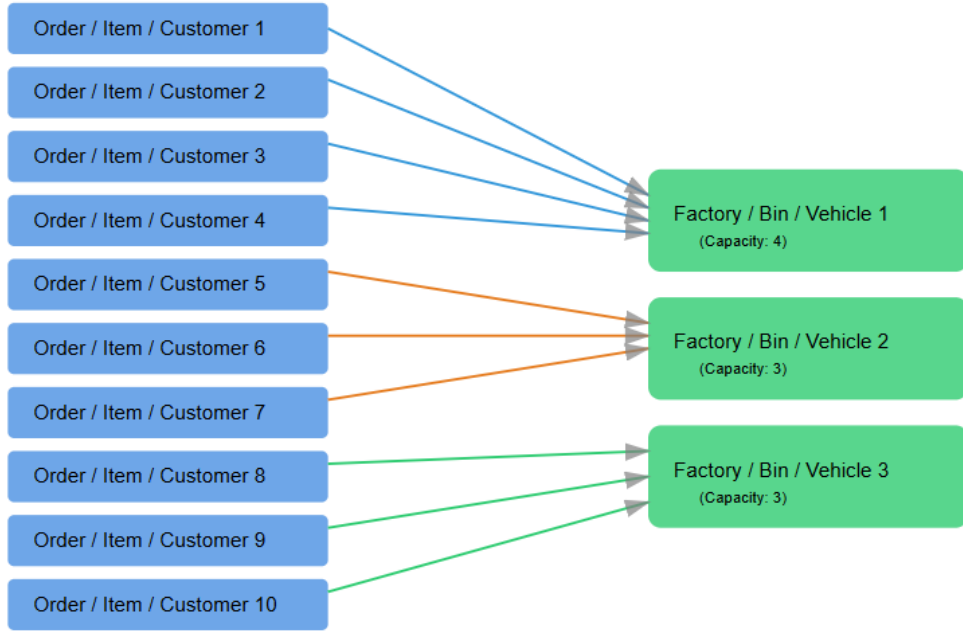


Figure 6: Similarities between BAP, BPP, and CVRP

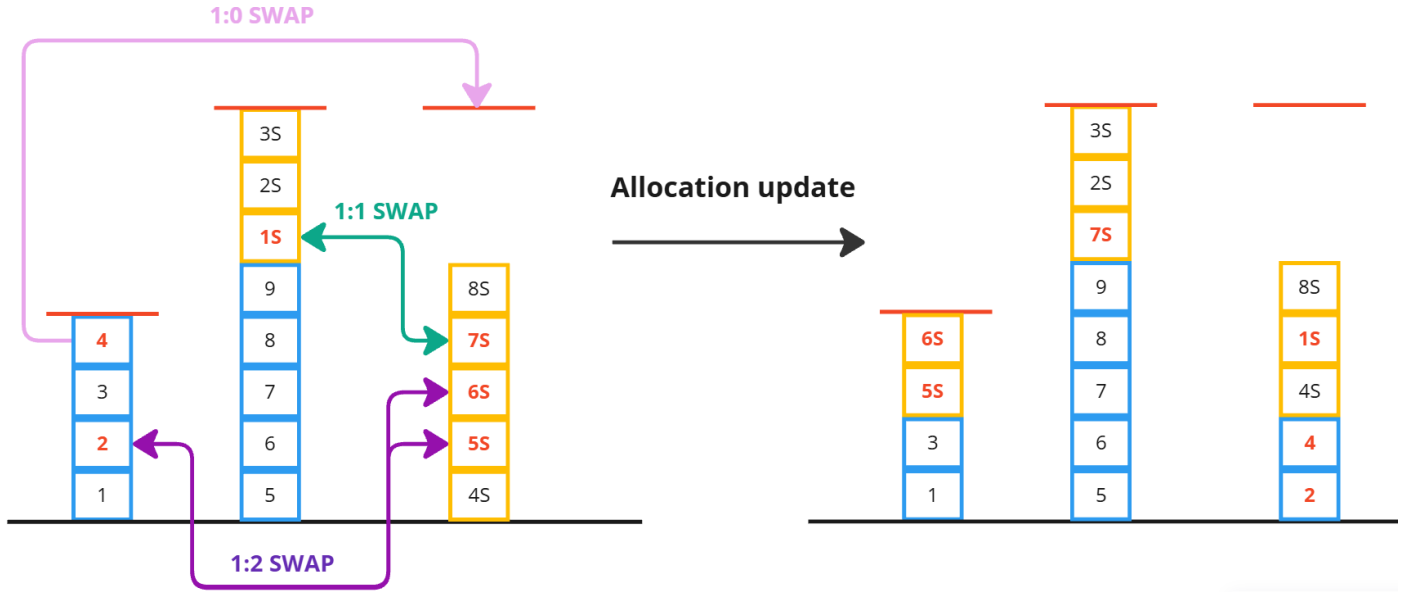


Figure 7: Types of box swaps

This article first examines the application of the exact method to address BAP. The CBC solver, developed under the Computational Infrastructure for Operations Research (COIN-OR) project, offers a robust open-source solution for MILP problems [6]. Its seamless integration with the COIN-OR linear programming solver and cut-generation library makes it particularly effective for tackling complex optimization challenges. In this study, the CBC solver can solve BAP in just a few steps. For example, consider a hypothetical order quantity of 100,000; the process proceeds as follows:

1. The model is loaded and initialized with 210,302 rows, 300,600 columns, and 1,359,124 elements. Pre-processing fixes 110,000 variables, reducing the problem size. Solving the continuous relaxation yields an objective value of 4,646.
2. The Feasibility Pump heuristic (Table 4) is employed to quickly obtain an initial integer-feasible solution. Coefficient Diving (Table 4) is then applied, successfully identifying the optimal solution. For most order quantities below 100,000, the Feasibility Pump alone suffices to find the optimal solution immediately.
3. A limited B&B process resolves any remaining general integer variables not fixed by heuristics. This step focuses on the reduced problem size, branching only on the unfixed variables.

4. A mini B&B is conducted to further refine the solution. While this step does not improve the objective, it serves to validate the quality of the heuristic-based solution.
5. The process concludes without extensive branching, as the optimal solution is proven at the root node following the limited B&B and mini B&B steps.

Table 4: Primal heuristics in CBC solver

Category	Feasibility Pump	Coefficient Diving
Overview	Introduced by Fischetti et al. [5], this heuristic rapidly finds feasible solutions by alternating between linear programming (LP) relaxation and integer solutions.	Based on works by Berthold [2], and Paulus and Krause [20], this heuristic prioritizes variables with minimal constraint violations.
Steps	<ol style="list-style-type: none"> 1. Start with the LP relaxation solution. 2. Round to the nearest integer-feasible point. 3. Find the closest LP solution to this integer point. 4. Repeat until a feasible solution is found or a stopping criterion is met. 	<ol style="list-style-type: none"> 1. Select variables based on <i>locks</i> (potential constraint violations). 2. Prioritize variables with minimal locks and smallest fractionality. 3. Iteratively bound selected variables to guide the solution toward feasibility.
Key advantage	The <i>pumping</i> action between LP-feasible and integer-feasible solutions enables fast convergence to high-quality feasible solutions.	This targeted approach efficiently finds optimal or near-optimal solutions early in the process.

Regarding the two heuristics, ITPS is inspired by the 2-Opt method commonly applied in VRP [25]. While 2-Opt swaps adjacent elements in a route to minimize travel distance or cost, ITPS swaps two orders between factories to minimize site-level WMAPE. ITPS has three defining features: it accepts only improving swaps, selects candidates based on recipe differences, and preserves feasibility by swapping only two orders at a time.

TS is a metaheuristic grounded in adaptive memory techniques, employing a memory structure known as the *tabu list* to guide the search more effectively [12]. The term *tabu* originates from Polynesian cultures, where it denotes prohibitions [21]. In TS, tabu moves are those temporarily forbidden to prevent cycling and encourage exploration of new regions in the solution space [11]. A central concept is the *tenure*, which defines the duration a move remains in the tabu list [9]. This mechanism prevents short-term cycling while allowing previously explored moves to be revisited after adequate diversification [1]. By balancing intensification (selecting the best move) with diversification (introducing random swaps), TS ensures a comprehensive search of the solution space.

4 Numerical experiments

4.1 Experimental design

Simulated data replicating the characteristics of real data is generated to evaluate the effectiveness of three proposed methods. The resulting order list follows realistic rules: 30% of the total orders are eligible for F1, 60% for F2, and all orders are eligible for F3, which has no constraints. When the proportion of real orders increases over time, all real orders from the previous day are carried over to the current day with their IDs unchanged, facilitating allocation tracking. Simulated orders are generated anew each day to make up for any shortfall in the total quantity, resulting in dynamic IDs. Each order consists of 1 to 4 recipes chosen from a total of 100 available recipes, with eligibility based on their grouping. Recipes in Group 1 (1–29) are eligible exclusively for F1, Group 2 (30–49) qualifies for both F1 and F2, Group 3 (50–89) is eligible only for F2, and Group 4 (90–100) is eligible solely for F3. An order qualifies for a factory only if all its recipes meet the factory’s eligibility. However, for F3-only orders, any recipe can be included as long as the order contains at least one recipe from Group 4. Table 5 presents exemplary order data along with solutions that satisfy the capacity constraints: 25% of the total quantity for F1 and 50% for F2. Importantly, the allocated factories for four existing real orders on LD12 remain unchanged, ensuring a minimum WMAPE site.

Table 5: Exemplary orders and solutions

LD12 (46% real orders)				LD11 (52% real orders)			
Order ID	Recipe IDs	Is real	Eligible factories	Order ID	Recipe IDs	Is real	Eligible factories
1	30	True	F1, F2, F3	1	30	True	F1, F2, F3
2	8, 5, 24	True	F1, F3	2	8, 5, 24	True	F1, F3
3	22	True	F1, F3	3	22	True	F1, F3
4	87	True	F2, F3	4	87	True	F2, F3
5	52, 51, 55, 63	False	F2, F3	5	74	True	F2, F3
6	82, 88	False	F2, F3	6	85	False	F2, F3
7	85	False	F2, F3	7	89, 73, 86	False	F2, F3
8	84, 76	False	F2, F3	8	54, 52	False	F2, F3
9	93, 1, 36, 76	False	F3	9	100, 99	False	F3
10	28, 95, 20	False	F3	10	91, 13	False	F3

Solution for LD12		Solution for LD11	
Factory	Allocated orders	Factory	Allocated orders
F1	2, 3	F1	2, 3
F2	1, 4, 5, 6, 8	F2	1, 4, 5, 7, 8
F3	7, 9, 10	F3	6, 9, 10

To evaluate the performance of the proposed methods, three key indicators are used. First, the optimality gap measures the difference between the MAPE site and global, with WMAPE global serving as the lower bound. CBC solver is expected to match WMAPE site to WMAPE global, as the exact method aims to find the optimal solution. Second, the improvement percentage measures the reduction in WMAPE site relative to the initial solution for heuristics, enabling performance comparison since these methods may not always reach the optimal result. Finally, optimization time is assessed by the duration each method takes to find the best solution, with a maximum limit of 10 minutes (600 seconds).

4.2 Results

4.2.1 Benchmark test

The following results are based on 10,000 orders over two consecutive days: LD12 (46% real orders) and LD11 (52% real orders). The CBC solver uses the LD12 allocation decision to directly determine the best allocation for LD11, minimizing the site-level WMAPE between the two days without swaps. For the heuristics, the number of iterations strongly influences their ability to reach high-quality solutions within reasonable time limits. To identify suitable iteration counts, tests were conducted on 30 different sets of 10,000 orders, resulting in 1,500 iterations for ITPS and 500 iterations for TS as the optimal settings. Using these values, the allocations for LD11 before and after applying ITPS and TS are summarized in Table 6. The results show that B&B or CBC solver remains the most effective method, achieving the optimal solution in the shortest time. Between the two heuristics, TS slightly outperforms ITPS in both solution quality and computational speed.

4.2.2 Scalability test

In this test, each method is applied to five different order quantities, ranging from 10,000 to 100,000, over two days: LD12 and LD11. Figure 8 presents the optimization times of the three algorithms, while Figures 9, 10, and 11 show the gap between site-level WMAPE and global WMAPE for each method. The results confirm that B&B consistently outperforms the heuristics, delivering optimal solutions across all quantities in the shortest time. An important observation is that both errors decrease as the total order quantity increases. This finding aligns with the principles of aggregate forecasting, which suggest that forecasts are more accurate for groups of items than for individual items, as variability within a group tends to cancel out, producing more stable and reliable predictions [22]. Moreover, with larger order quantities, individual fluctuations exert a smaller influence on overall error, whereas in smaller samples, changes in each order have a proportionally greater effect.

4.2.3 Temporal fixed test

In this test, the best-performing method is applied to address the temporal aspect of BAP. Specifically, B&B is used to make allocation decisions continuously over a 15-day planning horizon under ideal conditions, assuming no changes in factory capacities or actual orders, with a fixed daily order quantity of 10,000. Figure 12 illustrates the growing proportion of real orders across days. As LD3 approaches, the order composition between

Table 6: Benchmark test results

B&B	
LD12	
F1	2500 orders, 2500 real
F2	5000 orders, 1890 real
F3	2500 orders, 210 real
LD11	
F1	2500 orders, 2500 real
F2	5000 orders, 2593 real
F3	2500 orders, 107 real
WMAPE site	0.054
WMAPE global	0.054
Optimization time	2.89 seconds

	ITPS	TS
LD12		
F1	2500 orders, 2500 real	
F2	5000 orders, 1890 real	
F3	2500 orders, 210 real	
LD11 (Before)		
F1	2500 orders, 2500 real	
F2	5000 orders, 2422 real	
F3	2500 orders, 278 real	
LD11 (After)		
F1	2500 orders, 2500 real	2500 orders, 2500 real
F2	5000 orders, 2436 real	5000 orders, 2429 real
F3	2500 orders, 264 real	2500 orders, 271 real
WMAPE site		
Before	0.074	0.074
After	0.054	0.054
Improvement	26.21%	26.97%
WMAPE global		
0.054		
Optimization time	19.45 seconds	15.35 seconds

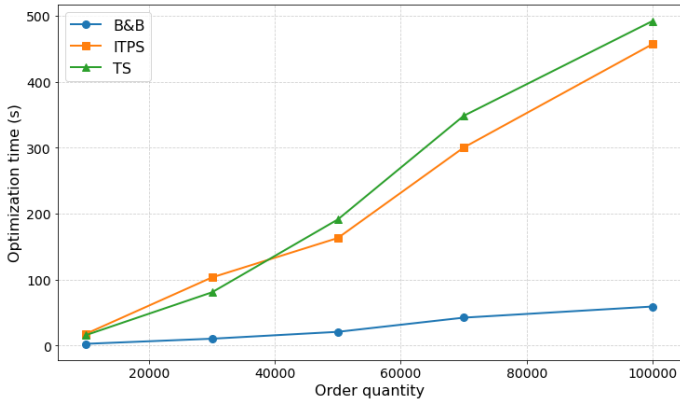


Figure 8: Comparison of optimization time

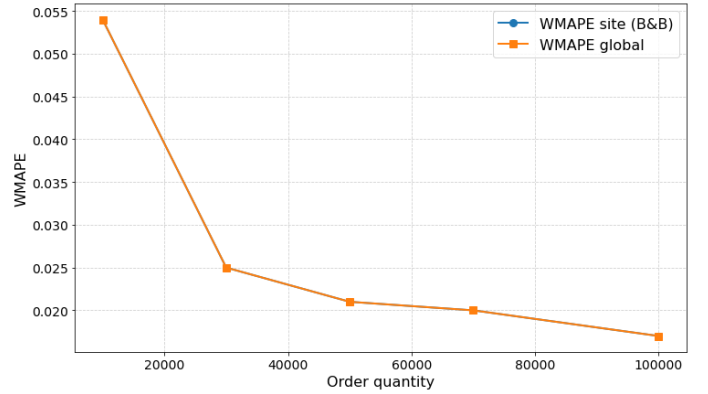


Figure 9: WMAPEs of B&B

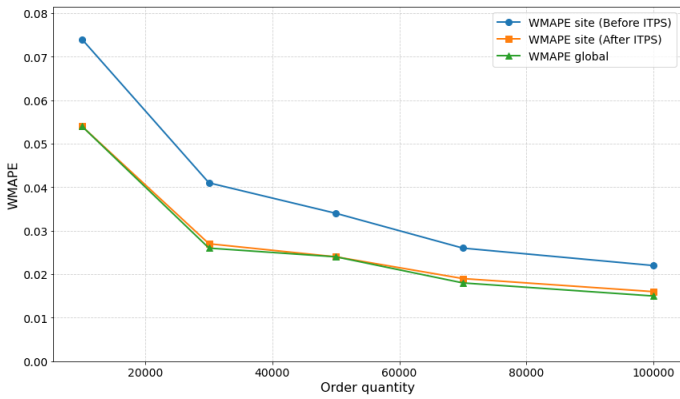


Figure 10: WMAPEs of ITPS

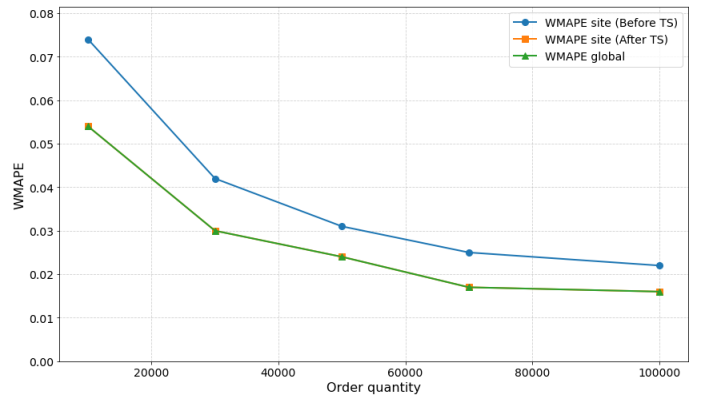


Figure 11: WMAPEs of TS

consecutive days becomes increasingly similar, since each day incorporates real orders carried forward from the previous day. This explains the observed decreasing trend in the error graphs.

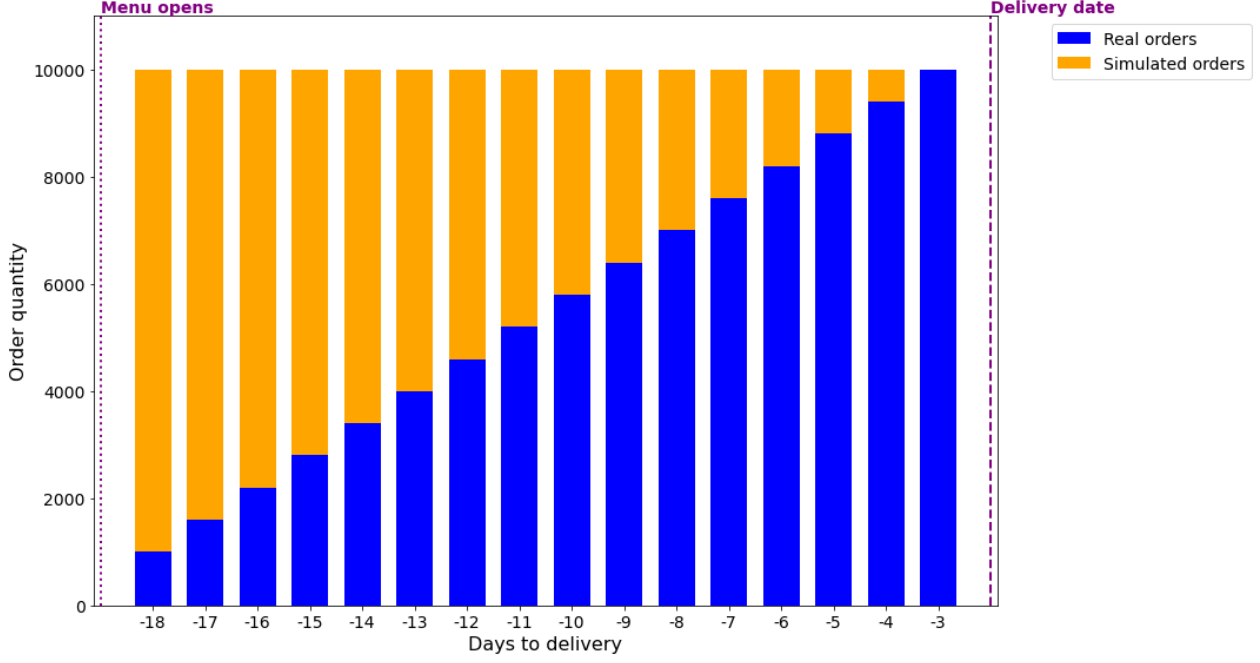


Figure 12: Composition of total orders through days

Figure 13 shows WMAPE site when B&B continuously optimizes allocation between two consecutive days. For instance, LD16 is optimized based on the B&B allocation of LD17, LD15 is optimized based on LD16’s allocation, and so on. It is evident that, except for some initial days when the proportion of real orders is still low, B&B consistently provides optimal solutions from LD14 to LD3. The WMAPE site (Greedy) line illustrates that, without B&B, the recipe differences between two consecutive days are significantly higher.

After B&B completes the order allocation for all days, each day’s decision is compared with the final day’s allocation, as shown in Figure 14. This comparison assesses whether the day-by-day optimization provides a smooth transition from soft allocations to hard allocation. The gradual decrease in WMAPE site indicates that B&B has effectively contributed to stable production planning, minimizing abrupt changes in the allocated recipe quantity. The ultimate goal is to reduce the area under the WMAPE site curve, ideally aligning it with the area under the WMAPE global curve. Although there is a big gap between the WMAPE global and WMAPE site (B&B) lines in the early days, this gap narrows over time. As a result, the area under the WMAPE site (B&B) curve is smaller than that under the WMAPE site (Greedy) curve, showing B&B’s effectiveness in achieving proxy optimization.

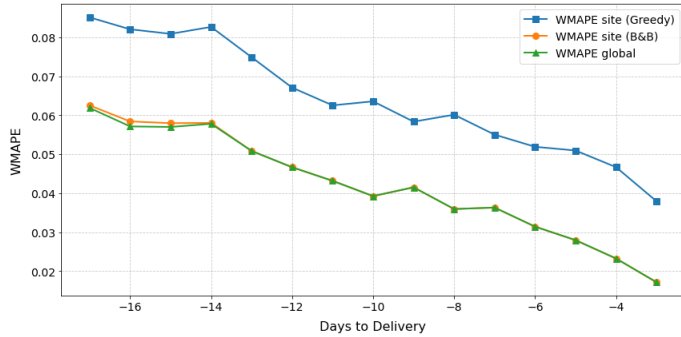


Figure 13: WMAPE site and WMAPE global through days

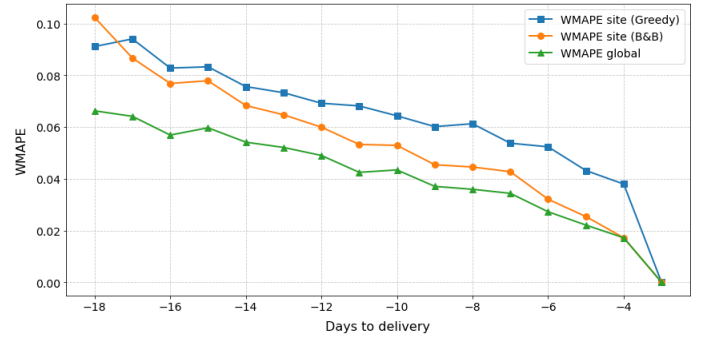


Figure 14: Each day’s allocation versus LD3’s allocation

4.2.4 Temporal variation test

In the previous test, factory capacities and real orders were assumed to remain constant throughout the entire planning period. However, in practice, the following scenarios may occur:

1. **Capacity change:** When F1's capacity drops from 3,000 to 1,000 orders on LD10, there is a sudden increase in WMAPE site, as shown in Figure 15. This spike occurs because orders must be rapidly redistributed among factories. B&B handles this effectively, quickly adjusting allocations and achieving the optimal WMAPE site by LD9. In contrast, the greedy method exhibits a large gap between WMAPE site and global after the capacity change, highlighting its lack of adaptability. Comparing each day's allocations with the final day's allocation in Figure 16 shows that the WMAPE site under B&B gradually approaches the optimal value, particularly from LD7 onward. The sharp drop in error between LD10 and LD11 reflects a significant shift in allocation decisions, with pre-LD10 allocations differing markedly from LD3. This test demonstrates B&B's ability to effectively respond to major changes in factory capacity.

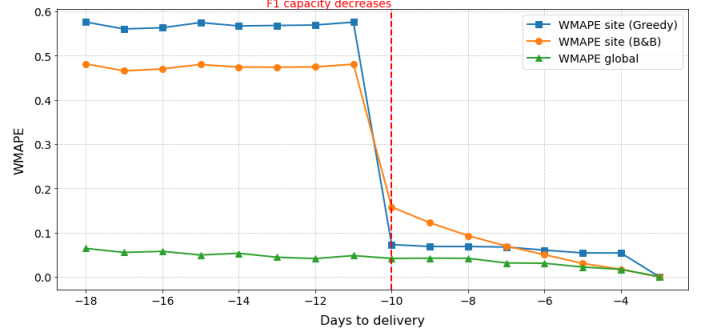
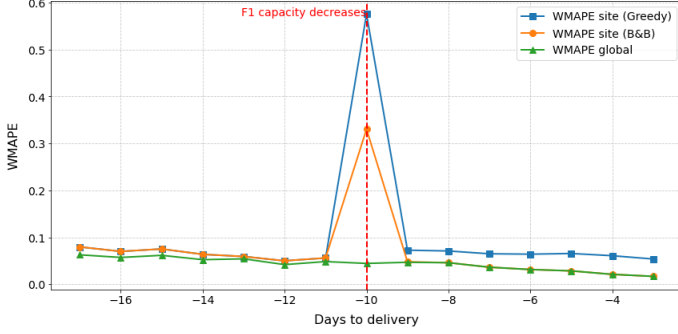


Figure 15: WMAPE site and WMAPE global through days

Figure 16: Each day's allocation versus LD3's allocation

2. **Order changes:** This study assumes that each day, customers delete 5% of real orders and modify recipes in 30% of them. Figure 17 illustrates how these changes induce fluctuations in WMAPEs. Without such modifications, both WMAPE site and global exhibit a clear downward trend. When changes occur, however, their values fluctuate considerably depending on the extent of adjustments. Despite this variability, the results confirm that B&B consistently achieves the optimal WMAPE site in all scenarios.

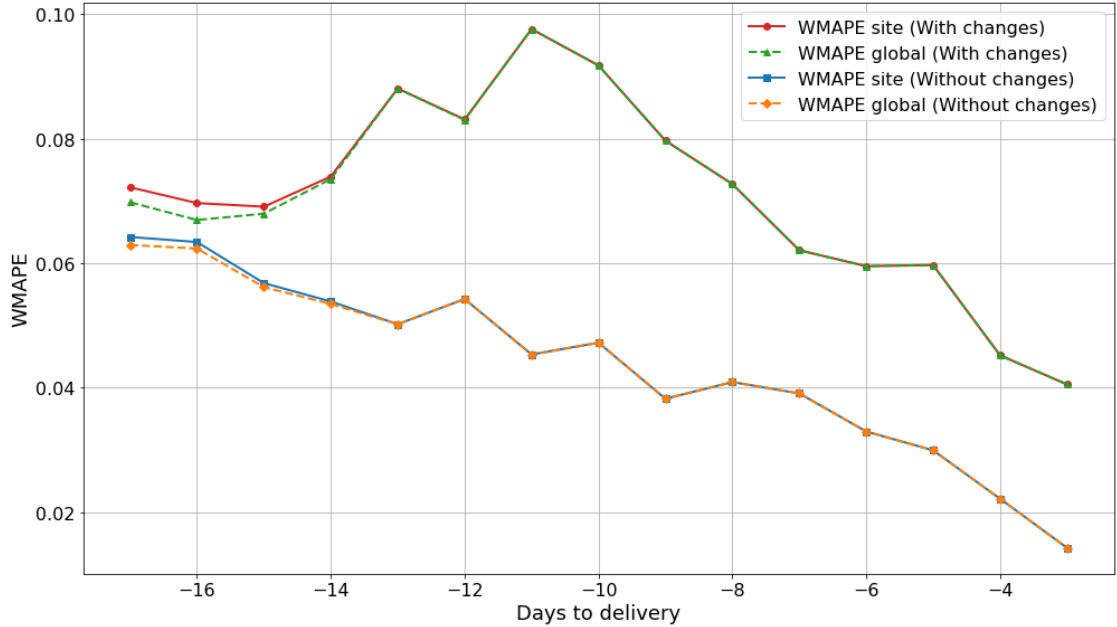


Figure 17: WMAPE site and WMAPE global with and without order changes

Also in this scenario, a simple allocation site strategy can be applied since the IDs of real orders remain unchanged across days. For example, if order O1 is assigned to F1 and O2 to F2 on LD14, the same assignments can be carried over to LD15. This ID-based approach ensures consistency and minimizes allocation changes for recurring orders. However, challenges emerge when recipes within an order change, potentially making it ineligible for its original factory. Deleted orders add further complications, as past real order IDs may no longer exist. Figures 18 and 19 illustrate these challenges by comparing the ID-based method with B&B. The results show that B&B consistently delivers optimal solutions, aligning WMAPE site with WMAPE global regardless of order changes. In contrast, even under stable condition,

the ID-based method aligns with WMAPE global only during later stages, when the proportion of real orders is high. Once order changes occur, the gap between WMAPE site (ID-based) and WMAPE global widens, underscoring the limitations of simple allocation method and the advantage of more advanced techniques like CBC solver.

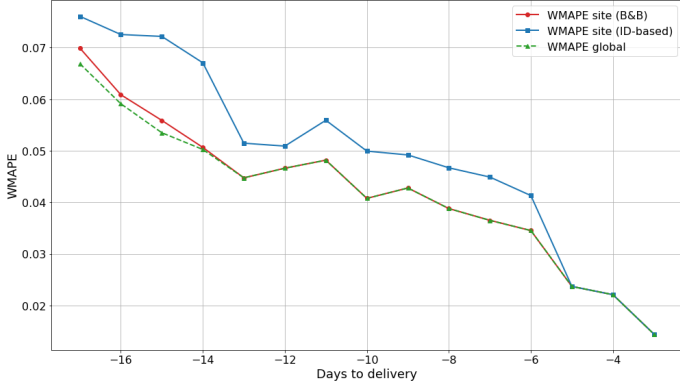


Figure 18: B&B and ID-based allocation without order changes

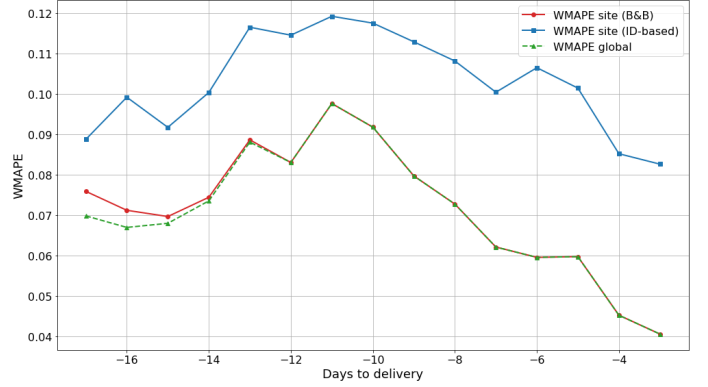


Figure 19: B&B and ID-based allocation with order changes

5 Conclusion

This study introduces BAP, a novel optimization challenge in meal kit delivery operations, offering both theoretical and practical insights for its resolution. Leveraging structural similarities with BPP and CVRP, we developed a MILP formulation and systematically evaluated exact versus heuristic solution approaches. Computational experiments show that CBC consistently delivers optimal solutions with remarkable efficiency and robustness under dynamic conditions, outperforming heuristic methods such as ITPS and TS.

Three significant findings emerge from this research. First, the inverse relationship between problem scale and allocation error indicates that consolidating allocation cycles can enhance operational performance. Second, the temporal dimension of BAP imposes unique constraints not captured by static models, underscoring the importance of time-consistent allocation strategies. Third, CBC’s superior performance in dynamic settings confirms the value of advanced optimization techniques over simple rule-based methods for practical implementation.

This study is limited by its focus on consecutive-day allocations. Future work could extend BAP to multi-day horizons, integrate demand uncertainty, and explore metaheuristics such as Genetic Algorithms or Large Neighborhood Search for richer neighborhood structure. Incorporating real-time data and collaborative supplier mechanisms could further improve adaptability and long-term strategic planning, ultimately enhancing efficiency, reducing food waste, and supporting sustainability objectives in meal kit delivery system.

ACKNOWLEDGMENTS

All numerical experiments were implemented in Python and performed on a personal computer with an Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz and 8 GB of RAM. This research received financial support from the University of Southampton. The funding source had no involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

DATA AVAILABILITY

The data that support the findings of this study are openly available in GitHub at https://github.com/35436506/recipe_box_allocation_optimization_in_meal_kit_delivery.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

ORCID

Thi Minh Thu Nguyen <https://orcid.org/0009-0000-9058-3913>

Alain Zemkoho <https://orcid.org/0000-0003-1265-4178>

REFERENCES

- [1] Battiti, R. and Tecchiolli, G. (1994) 'The reactive tabu search', *ORSA Journal on Computing*, 6(2), pp. 126-140. Available at: <https://doi.org/10.1287/ijoc.6.2.126>.
- [2] Berthold, T. (2006) *Primal heuristics for mixed integer programs*. Diploma thesis. Technische Universität Berlin. Available at: https://www.researchgate.net/publication/258846101_Primal_Heuristics_for_Mixed_Integer_Programs.
- [3] Clausen, J. (1999) 'Branch and bound algorithms-principles and examples', *Department of Computer Science, University of Copenhagen*, pp. 1-30. Available at: <https://imada.sdu.dk/~jbj/DM85/TSPtext.pdf>.
- [4] Clear Spider (2024) *How Meal Kit Delivery Services Manage Their Supply Chains*. Available at: <https://clearspider.net/blog/meal-kit-delivery-supply-chains/>.
- [5] Fischetti, M., Glover, F. and Lodi, A. (2005) 'The feasibility pump', *Mathematical Programming*, 104(1), pp. 91-104. Available at: <https://doi.org/10.1007/s10107-004-0570-3>.
- [6] Forrest, J. and Lougee-Heimer, R. (2005) 'CBC user guide', in *INFORMS Tutorials in Operations Research*, pp. 257-277. Available at: <https://doi.org/10.1287/educ.1053.0020>.
- [7] GeeksforGeeks (2021) *How to Calculate Weighted MAPE in Excel?*. Available at: <https://www.geeksforgeeks.org/how-to-calculate-weighted-mape-in-excel/>.
- [8] Gendreau, M., Hertz, A. and Laporte, G. (1994) 'A Tabu Search Heuristic for the Vehicle Routing Problem', *Management Science*, 40(10), pp. 1276-1290. Available at: <https://doi.org/10.1287/mnsc.40.10.1276>.
- [9] Gendreau, M. (2003) 'An introduction to tabu search', in Glover, F. and Kochenberger, G.A. (eds.) *Handbook of metaheuristics*. Boston: Springer, pp. 37-54. Available at: https://doi.org/10.1007/0-306-48056-5_2.
- [10] Glover, F. (1986) 'Future paths for integer programming and links to artificial intelligence', *Computers & Operations Research*, 13(5), pp. 533-549. Available at: [https://doi.org/10.1016/0305-0548\(86\)90048-1](https://doi.org/10.1016/0305-0548(86)90048-1).
- [11] Glover, F. (1989) 'Tabu search—part I', *ORSA Journal on Computing*, 1(3), pp. 190-206. Available at: <https://courses.cs.umass.edu/cics521-cg/docs/tabu-glover-1.pdf>.
- [12] Glover, F. and Laguna, M. (1997) *Tabu search*. Boston: Springer Science & Business Media. Available at: <http://dx.doi.org/10.1007/978-1-4615-6089-0>.
- [13] Langham Logistics (2020) *The Complex Logistics Behind Meal-Kit Services*. Available at: <https://www.elangham.com/whats-for-dinner-the-complex-logistics-behind-meal-kit-services/>.
- [14] Laporte, G. and Nobert, Y. (1987) 'Exact algorithms for the vehicle routing problem', *Annals of Discrete Mathematics*, 132, pp. 147-184. Available at: [https://doi.org/10.1016/S0304-0208\(08\)73235-3](https://doi.org/10.1016/S0304-0208(08)73235-3).
- [15] Lin, S. (1965) 'Computer solutions of the traveling salesman problem', *Bell System Technical Journal*, 44(10), pp. 2245-2269. Available at: <https://doi.org/10.1002/j.1538-7305.1965.tb04146.x>.
- [16] Loic, G. (2024) *Internal material*. Internal company document.
- [17] Lodi, A., Martello, S. and Vigo, D. (1999) 'Heuristic and metaheuristic approaches for a class of two-dimensional bin packing problems', *INFORMS Journal on Computing*, 11(4), pp. 345-357. Available at: <https://doi.org/10.1287/ijoc.11.4.345>.
- [18] Martello, S. and Toth, P. (1990) *Knapsack Problems: Algorithms and Computer Implementations*. Wiley. Available at: http://old.math.nsc.ru/LBRT/k5/knapsack_problems.pdf.

- [19] Osman, I.H. (1993) 'Metastrategy simulated annealing and tabu search algorithms for the vehicle routing problem', *Annals of Operations Research*, 41(4), pp. 421-451. Available at: <https://doi.org/10.1007/BF02023004>.
- [20] Paulus, M.B. and Krause, A. (2023) 'Learning To Dive In Branch And Bound', *arXiv preprint arXiv:2301.09943*. Available at: <https://doi.org/10.48550/arXiv.2301.09943>.
- [21] Radcliffe-Brown, A.R. (1939) *Taboo*. Cambridge: Cambridge University Press.
- [22] Reid, R.D. and Sanders, N.R. (2012) *Operations Management: An Integrated Approach*. 5th edn. Wiley.
- [23] RMS Omega (2024) *4 Challenges For Meal Kit Delivery Operations*. Available at: <https://rmsomega.com/4-challenges-for-meal-kit-delivery-operations/>.
- [24] Statista (2024) *Meal Kits in the UK - Overview*. Available at: <https://www.statista.com/topics/8895/meal-kits-in-the-uk/>.
- [25] Toth, P. and Vigo, D. (eds.) (2002) *The Vehicle Routing Problem*. SIAM Monographs on Discrete Mathematics and Applications. Available at: <https://api.semanticscholar.org/CorpusID:209742615>.
- [26] UN Climate Change News (2024) *Food loss and waste account for 8-10% of annual global greenhouse gas emissions; cost USD 1 trillion annually*. Available at: <https://unfccc.int/news/food-loss-and-waste-account-for-8-10-of-annual-global-greenhouse-gas-emissions-cost-usd-1-tril>
- [27] Woop (2024) *Home*. Available at: <https://woop.co.nz/>.
- [28] Zangl, G., Graf, T. and Al-Kinani, A. (2006) 'Proxy Modeling in Production optimization', Paper presented at the *SPE Europe/EAGE Annual Conference and Exhibition*, Vienna, Austria, 12-15 June. Available at: <https://doi.org/10.2118/100131-MS>.