Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

RESEARCH ARTICLE

Box Allocation Optimization in Meal Kit Delivery

Thu T. M. Nguyen[1] | Loïc Genest[2] | Alain Zemkoho[3]

[1]Department of Decision Analytics and Risk, Southampton Business School, Southampton, United Kingdom

[2]Senior Data Scientist, London, United Kingdom

[3]School of Mathematical Sciences, University of Southampton, Southampton, United Kingdom

Correspondence
Thu T. M. Nguyen, Department of Decision Analytics and Risk, Southampton Business School, Southampton, UK.
Email: thu.minh090997@gmail.com

Abstract

This study investigates the Box Allocation Problem (BAP), a novel and underexplored challenge in meal kit delivery. BAP aims to minimize recipe variations in daily allocation decisions while adhering to capacity and eligibility constraints. Addressing this gap, the research formulates a mathematical model and evaluates both exact and heuristic optimization methods. Results demonstrate the superior performance of the Branch and Bound algorithm in the CBC solver, achieving optimal solutions under dynamic conditions and outperforming heuristic approaches like Tabu Search and Iterative Targeted Pairwise Swap. By improving operational efficiency and promoting sustainability through reduced allocation errors, this study provides actionable insights for optimizing strategies in the growing meal kit delivery market.

KEYWORDS
Meal kit delivery, Box allocation problem, Branch and Bound, CBC solver, Heuristics

1 | INTRODUCTION

The meal kit delivery market in the UK has experienced significant growth in recent years, driven by increasing consumer demand for convenient, healthy, and personalized meal options. With a projected revenue of $1.4 billion in 2024, the industry has become highly competitive, attracting both established companies and new entrants [24]. Success in this sector depends heavily on efficient supply chain management, ensuring the timely delivery of fresh ingredients. However, the dynamic nature of the market presents several challenges, including inventory management, transportation efficiency, and the handling of perishable goods. These difficulties make logistics a critical factor in determining success [13,3]. Moreover, companies must quickly adapt to fluctuations in demand and potential disruptions, which makes flexible and responsive supply chain strategies essential for maintaining competitiveness [23].

One major challenge is the Box Allocation Problem (BAP), which involves assigning recipe boxes as Figure 1 to factories in a way that minimizes recipe variations between days, measured by the Weighted Mean Average Percentage Error (WMAPE), while adhering to capacity and eligibility constraints [17]. Minimizing daily WMAPE site (factory) is critical for two key reasons: first, it helps to reduce waste. By maintaining consistent recipe allocations across factories, we achieve more accurate ingredient forecasting. This leads to less over-ordering and reduced food spoilage, enhancing sustainability. Second, it provides better item availability for customers. Stable allocations help maintain consistent ingredient stocks across our factories, reducing the risk of stockouts. This results in more reliable order fulfillment, improving customer satisfaction. Despite its importance to operational efficiency, BAP remains unexplored in both academic and industry literature.

FIGURE 1 Recipe box [26]

This study aims to fill the research gap by developing a comprehensive mathematical model for BAP, integrating the unique constraints of factory capacity and recipe eligibility. Furthermore, the research systematically compares exact method, specifically Branch and Bound (B&B) in CBC solver, with heuristics, including 2-Opt or Iterative Targeted

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; TS, Tabu Search; WMAPE, Weighted Mean Absolute Percentage Error.

Pairwise Swap (ITPS) and Tabu Search (TS), which are commonly used for combinatorial optimization problems like the Bin Packing Problem (BPP) and Vehicle Routing Problem (VRP) [14,15,8]. These heuristic methods, especially TS, are known for their effectiveness in navigating large solution spaces and overcoming local optima [19,16], while B&B ensures optimality through systematic exploration and pruning of infeasible solutions [18,4]. This study evaluates the performance of these approaches across different order quantities and operational scenarios, providing insights into their scalability and adaptability.

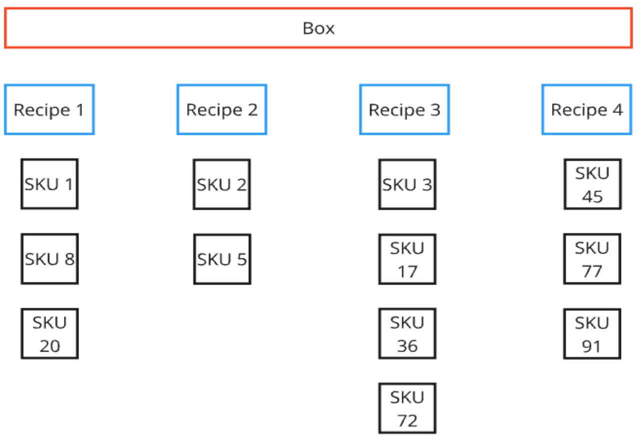
The results highlight the superiority of CBC solver in achieving optimal solutions, even in dynamic environments with fluctuating factory capacities and orders. By addressing both theoretical and practical aspects of BAP, this research contributes to the literature on supply chain optimization and offers actionable insights for meal kit delivery companies. It bridges the gap between existing knowledge and the industry's need for scalable solutions that improve operational efficiency and sustainability.

The remainder of this paper is structured as follows: Section 2 details the characteristics of BAP and its mathematical formulation, comparing it with two common problems. Section 3 outlines the optimization methods, including both exact and heuristic approaches. Section 4 presents the results of numerical experiments, comparing the performance of these methods. Finally, Section 5 discusses the implications of the findings, identifies limitations, and proposes directions for future research.

2 | PROBLEM STATEMENT

2.1 | Problem background

BAP in meal kit delivery involves assigning boxes with structure as Figure 2a to different factories as Figure 2b. The goal is to minimize the recipe differences between two allocation decisions. For instance, if Factory 1 (F1) was assigned 100 units of Recipe 1 on the previous day, the company will allocate current orders so that F1 continues to produce 100 units of Recipe 1.



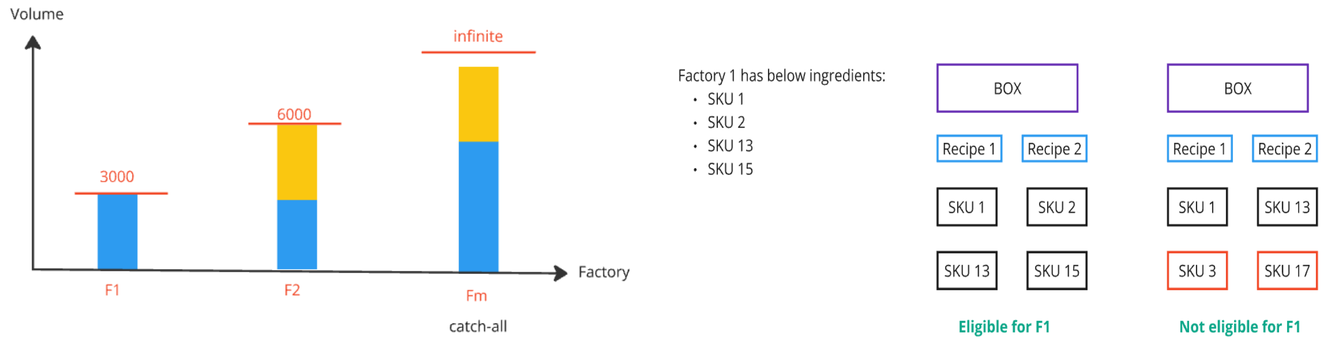
(a) Box structure [17]

(b) Allocation process [17]

FIGURE 2 Recipe box allocation

Regarding the first constraint, each factory, except the catch-all factory has a daily capacity, defined by the number of boxes it can produce. To ensure maximum resource utilization, the solution must fully meet this capacity as shown in Figure 3a. Additionally, because not all factories can produce every recipe, the allocation decision must account for the eligibility constraint explained in Figure 3b, ensuring that each box is only assigned to the factory capable of processing its recipes.

A key feature of BAP is its temporal nature, requiring the daily, or even twice-daily, allocation of both simulated and real orders, as demonstrated in Figure 4. In this study, a 15-day planning period is assumed, starting from lead day 18 (LD18) and ending on LD3. The planning process is as follows: initially, the company forecasts the total number of boxes expected to be ordered by customers for the upcoming time. Throughout the planning period, the company monitors actual orders against the projection. When real orders are lower than the total forecasted quantity, the ordering platform is used to generate simulated orders to represent the anticipated future demand. The company then performs *soft allocation*, distributing both real and simulated orders across factories. As time progresses, the proportion of real orders in the total quantity increases, gradually replacing the simulated orders. Importantly, the allocation for each day includes all real orders from the previous day. On the final day (LD3), the allocation consists entirely of real orders. This stage, known as *hard allocation*, is fixed and no longer subject to changes.



(a) Capacity constraint [17]

(b) Eligibility constraint [17]

FIGURE 3 Main constraints

Then, the company is assumed to have three days to produce and deliver the boxes to customers. This approach enables the company to maintain consistent daily allocations while accommodating the dynamic nature of incoming orders throughout the planning period. It facilitates a structured transition from forecast-based planning to actual order fulfillment.

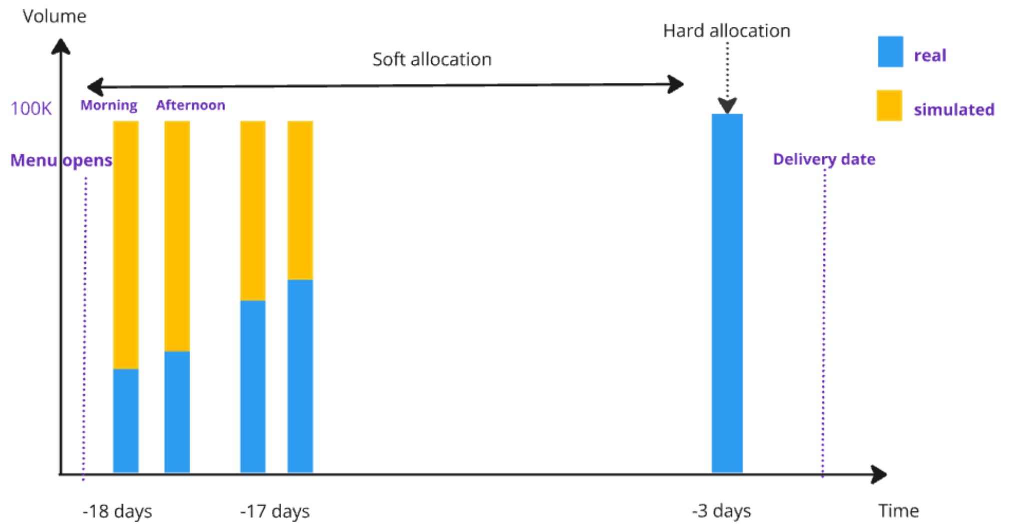


FIGURE 4 Temporal aspect of BAP [17]

2.2 | Mathematical formulation

Weighted Mean Absolute Percentage Error (WMAPE) is an advanced forecast accuracy measure in business contexts like supply chain and production planning. It sums absolute differences between actual and forecasted values, divided by the sum of actual values [7]. A lower WMAPE indicates better accuracy. Based on the characteristics of WMAPE site and WMAPE global shown in Table 1, this study focuses on minimizing WMAPE site which offers a detailed perspective and allows planners to assess the consistency of production at 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 following notations are used in the formulation of BAP:

t: Allocation day
$N = \{1, 2, \ldots, n\}$: Set of recipes, indexed by i
$S = \{1, 2, \ldots, m\}$: Set of factories, indexed by j
$O = \{1, 2, \ldots, k\}$: Set of orders, indexed by o
$a_{i,j,t}$: Number of times recipe i is allocated to factory j at day t
$a_{i,j,t-1}$: Number of times recipe i is allocated to factory j at day $t-1$
$a_{i,t-1}$: Number of recipe i at day $t-1$
$a_{i,t}$: Number of recipe i at day t
$x_{o,j,t}$: Binary decision variable indicating the allocation of order o to factory j at day t
$C_{j,t}$: Capacity of factory j at day t
$E_{i,j,t}$: Binary parameter indicating the eligibility of recipe i at factory j at day t

BAP can be formulated as a mixed-integer linear programming problem:

$$\underset{x}{\text{Minimize}}\, f(x) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} \left|a_{i,j,t} - a_{i,j,t-1}\right|}{\sum_{i=1}^{n} a_{i,j,t}} \tag{2.1}$$

subject to

$$\sum_{o=1}^{k} x_{o,j,t} = C_{j,t}, \quad \forall j \in S \setminus \{m\}, \forall t \tag{2.2}$$

$$x_{o,j,t} \le \min_{i \in o} E_{i,j,t}, \quad \forall o \in O, \forall j \in S, \forall t \tag{2.3}$$

$$\sum_{j=1}^{m} x_{o,j,t} = 1, \quad \forall o \in O, \forall t \tag{2.4}$$

$$x_{o,j,t} \in \{0, 1\}, \quad \forall o \in O, \forall j \in S, \forall t \tag{2.5}$$

The objective function (2.1) aims to minimize the WMAPE site for recipe allocations across all factories. Specifically, it reduces the day-to-day fluctuation in the number of recipes allocated to each factory while considering the total number of recipes allocated on the current day. Although the objective function does not explicitly include the binary decision variable $x_{o,j,t}$, the number of recipes $a_{i,j,t}$ is derived from the allocation of orders through the constraints. The constraints ensure that the decision variables $x_{o,j,t}$ directly determine which orders are assigned to which factories, thereby influencing the resulting recipe counts at each factory.

Constraint (2.2) ensures the total number of orders allocated to each factory (excluding the catch-all factory F_m) matches the factory's capacity. Constraint (2.3) guarantees that an order o is assigned to a factory j only if all recipes in that order are eligible for production at that factory. Constraint (2.4) ensures each order o is assigned to exactly one factory. Constraint (2.5) restricts the allocation decision variables to binary values, representing whether or not an order is assigned to a factory.

WMAPE global calculated by Equation (2.6) serves as a benchmark for evaluating the optimality of allocation solution because it is the lower bound of WMAPE site explained by Equations (2.7) and (2.8).

$$WMAPE_{\text{global}} = \frac{\sum_{i=1}^{n} \left|a_{i,t} - a_{i,t-1}\right|}{\sum_{i=1}^{n} a_{i,t}} \tag{2.6}$$

Decomposing at the site level, equation (2.6) becomes:

$$WMAPE_{\text{global}} = \frac{\sum_{i=1}^{n} \left|\sum_{j=1}^{m} a_{i,j,t} - \sum_{j=1}^{m} a_{i,j,t-1}\right|}{\sum_{i=1}^{n} a_{i,t}} \tag{2.7}$$

Using the triangle inequality, we get:

$$WMAPE_{\text{global}} \le \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} \left|a_{i,j,t} - a_{i,j,t-1}\right|}{\sum_{i=1}^{n} a_{i,t}} = WMAPE_{\text{site}} \tag{2.8}$$

Table 2 presents exemplary order data for two consecutive days, demonstrating their allocation while Table 3 illustrates the calculation results of WMAPE site and global. It can be seen that despite using the same order data, WMAPE global (0.5) is lower than WMAPE site (1.14).

The company's ultimate goal is to minimize area under the curve in Figure 5, which represents WMAPE site when comparing the allocation of the final day with the allocation of all days. However, directly achieving this goal is not feasible due to uncontrollable forecast errors before LD3. These errors can arise from various factors, including the randomness of simulated orders, modifications made by customers to existing real orders, and the arrival of new real orders. To tackle this challenge, the

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** ($a_{i,j,t}$)	**Absolute difference**
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** ($a_{i,t}$)	**Absolute difference**
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			

company employs a more practical strategy: reducing errors between consecutive days as Equation 2.1. This turns BAP into a *proxy optimization problem*, where a simpler, related problem is solved in place of the more complex, ultimate objective[27]. In this context, minimizing daily errors serves as a proxy for reducing the total error over the entire period.

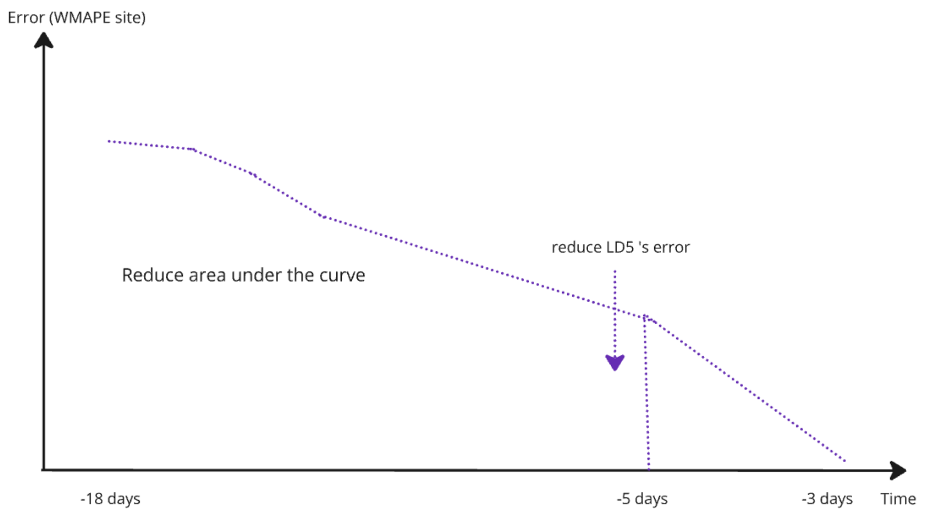


FIGURE 5 Objective of proxy optimization [17]

BAP shares similarities with two main classical problems: BPP and CVRP. These problems can be framed as network optimization tasks, where nodes represent entities such as factories, bins, or vehicles, and edges denote decisions like order allocations, item assignments, or routing paths, all subject to capacity constraints, as depicted in Figure 6. Despite these similarities, three problems have distinct mathematical formulations tailored to their respective contexts. For instance, BPP minimizes the number of bins required to pack items, ensuring that the total weight or volume in each bin does not exceed its capacity while CVRP minimizes travel costs or distances, ensuring that each vehicle adheres to its capacity. The similarities between BPP, CVRP, and the emerging BAP suggest they can share optimization techniques. For example, the B&B method, extensively used in BPP for solving capacity-constrained assignments, has been effectively applied in various allocation scenarios [18]. Similarly, heuristics like 2-Opt and TS, prominent in CVRP for refining route allocations by swapping customers between vehicles, have significantly influenced the development of strategies in allocation problems [15,10]. However, applying these techniques to BAP requires adaptations to account for its unique objective function and constraints, such as recipe eligibility and day-to-day consistency. These distinctions highlight how BAP extends beyond the traditional scope of BPP and CVRP, requiring problem-specific mathematical modeling while leveraging established methods to address its optimization challenges.

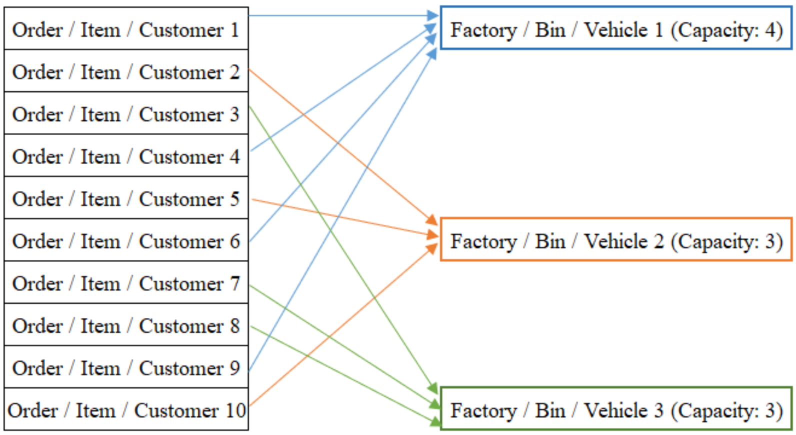


FIGURE 6 Exemplary solution of BAP, BPP, and CVRP

3 | SOLVING THE PROBLEM

The company seeks an effective method to minimize recipe disparities between two allocation solutions. In this study, the 1:1 box swap in Figure 7 is used for heuristics because swapping multiple orders at once risks creating ineligible moves or failing to meet full factory capacity. The solving process is as follows:

- For the exact method: directly apply CBC solver to current orders to determine their optimal allocation that minimizes the recipe difference with the previous day's allocation.
- For heuristics, the process begins by constructing an initial feasible solution using a greedy algorithm. Orders are allocated to F1 first, prioritizing orders with fewer eligible factories (F1-F3, then F1-F2-F3) until its capacity is reached. The remaining orders with eligible factories F1-F2-F3 and F2-F3 are allocated to F2 until its capacity is reached. All unallocated orders are assigned to F3, the catch-all factory with no constraints. Next, ITPS and TS are applied to improve the initial allocation.
- Performance and scalability of each method will be evaluated by comparing the final WMAPE site against WMAPE global and comparing optimization times across different quantities. This analysis will help identify the most effective method. The best one will be used to conduct temporal tests, assessing its effectiveness in optimizing allocations over a 15-day planning period and its ability to adapt to dynamic changes.

This article first explores the application of the exact method to address BAP. CBC solver, developed as part of 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 intricate optimization challenges. In this case, CBC solver can solve BAP with just a few steps. Take order quantity of 100,000 (the highest demand) for example. The process is as follows:

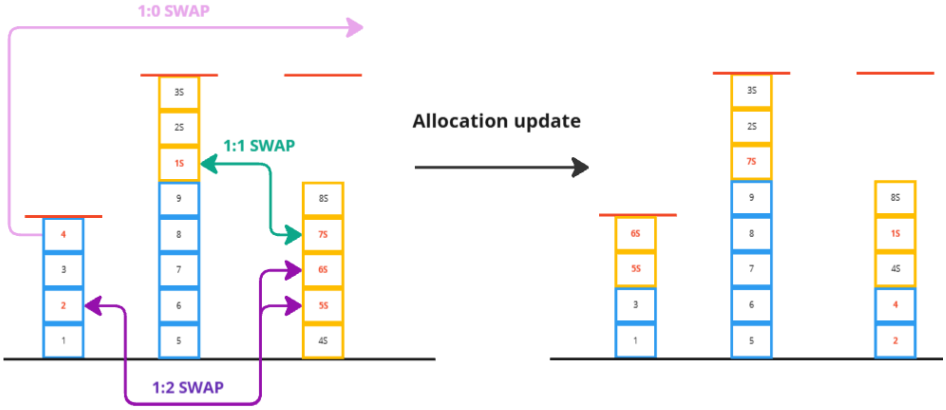


FIGURE 7 Exemplary types of box swaps

1. The model was loaded and initialized with 210,302 rows, 300,600 columns, and 1,359,124 elements. Preprocessing was applied, which fixed 110,000 variables and reduced the problem size. The continuous relaxation was solved, yielding an objective value of 4,646.
2. Employ Feasibility Pump heuristic explained in Table 4 to quickly find initial integer-feasible solutions. Then apply Coefficient Diving in Table 4, which successfully found the optimal solution. For most order quantities under 100,000, Feasibility Pump alone can find the optimal solution immediately.
3. A limited B&B process was used to clear up the remaining general integer variables that were not fixed by the heuristics. This process focused on a reduced problem size, branching only on the unfixed general integer variables.
4. A mini B&B was conducted to enhance the solution. Although this process ultimately did not lead to an improved solution, it plays a crucial role in validating the quality of the solution obtained through heuristics.
5. The process concluded without further extensive branching, as the optimal solution was proven at the root node after the limited branching and mini B&B processes.

TABLE 4 Primal heuristics in CBC

Feasibility Pump	**Coefficient Diving**
Introduced by Fischetti et al.[5], this heuristic quickly finds feasible solutions through an alternating process:	Based on works by Berthold[2], and Paulus and Krause[20], this heuristic focuses on variables with minimal constraint violations:
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.	1. Selects variables based on their *locks* (potential constraint violations). 2. Chooses variables with minimal locks and the smallest fractionality. 3. Iteratively bound selected variables to guide the solution toward feasibility.
The *pumping* action between LP-feasible and integer-feasible solutions rapidly converges to high-quality feasible solutions.	This targeted approach often finds the optimal or near-optimal solutions early in the process.

Regarding two heuristic methods, Iterative Targeted Pairwise Swap (ITPS) is inspired by local search, particularly the 2-Opt commonly used for VRP. While 2-Opt swaps adjacent elements in a tour, ITPS swaps two orders between two factories to reduce WMAPE site. ITPS is characterized by three key features: only accepting improvements, using recipe differences to select swap candidates, and maintaining solution feasibility by swapping only two orders at once.

Tabu Search (TS) is a metaheuristic inspired by adaptive memory techniques used in local search. TS distinguishes itself from local search by incorporating a memory structure, known as the *tabu list*, to guide the search process more intelligently[12]. The term *tabu* originates from Polynesian cultures, where it refers to prohibited things[21]. In TS, tabu moves are those that are temporarily forbidden to prevent cycling and encourage the exploration of new areas in the solution space[11]. A key concept in TS is *tenure*, which refers to the duration for which a move remains in the tabu list[9]. The tabu tenure determines how long a recently performed move is forbidden, avoiding short-term cycling and allowing the search to revisit previously explored areas after sufficient exploration elsewhere[1]. The balance between intensification (through selecting the best move) and diversification (through random swaps) provides a thorough exploration 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 adheres to realistic rules: 30% of total orders are eligible for F1, 60% for F2, and all orders are eligible for F3, which has no constraints. As 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, enabling seamless allocation tracking. Simulated orders, however, 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 (recipes 1–29) are eligible exclusively for F1. Group 2, which includes recipes 30–49, qualifies for both F1 and F2. Recipes in Group 3, spanning 50–89, are eligible only for F2. Lastly, Group 4 encompasses recipes 90–100 and is eligible solely for F3. An order qualifies for a factory only if all of 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 example order data along with solutions that satisfy the capacity constraints: 25% of the total quantity for F1 and 50% for F2. Notably, the allocated factories for the four existing real orders remain unchanged, ensuring a minimum WMAPE site.

TABLE 5 Exemplary orders and solutions

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

LD11 (52% real orders)			
Order ID	**Recipe IDs**	**Is real**	**Eligible factories**
1	30	True	F1, F2, F3
2	8, 5, 24	True	F1, F3
3	22	True	F1, F3
4	87	True	F2, F3
5	74	True	F2, F3
6	85	False	F2, F3
7	89, 73, 86	False	F2, F3
8	54, 52	False	F2, F3
9	100, 99	False	F3
10	91, 13	False	F3

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

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

To evaluate the performance of the proposed methods, three key indicators are used. 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 exact method can find the optimal solution. The improvement percentage calculates the percentage of WMAPE site improvement over the initial solution for heuristic methods. This helps compare the performance of different heuristics, as they may not always achieve the optimal result. Finally, computation time measures the time taken by each method to find the best solution, with a maximum allowed time of 10 minutes (600 seconds). All tests are implemented in Python on a PC with an Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz and 8 GB of RAM, and the codes are available on Github.

4.2 | Results

4.2.1 | Benchmark test

The following results are based on the total quantity of 10,000 orders over two consecutive days: LD12 (46% real orders) and LD11 (52% real orders). B&B uses the allocation decision of LD12 in Figure 8 to directly determine the best allocation for LD11, which minimizes WMAPE site between two days (no swaps like heuristics). The allocation result of B&B is shown in Figure 9 while more detail can be found in Table 6. For heuristics, the number of iterations significantly influences their ability to find the optimal solution within a reasonable time. Therefore, the iteration test is conducted on 30 different sets of 10,000 orders and identified that 1,500 and 500 is the best number of iterations for ITPS and TS respectively. With these iterations, the allocations for LD11 before and after applying ITPS and TS are obtained in Figure 11 and Figure 13. Table 6 points out that

B&B is the best method, as it provides optimal solution in the shortest time. Regarding heuristics, TS slightly outperforms ITPS in both solution quality and speed.

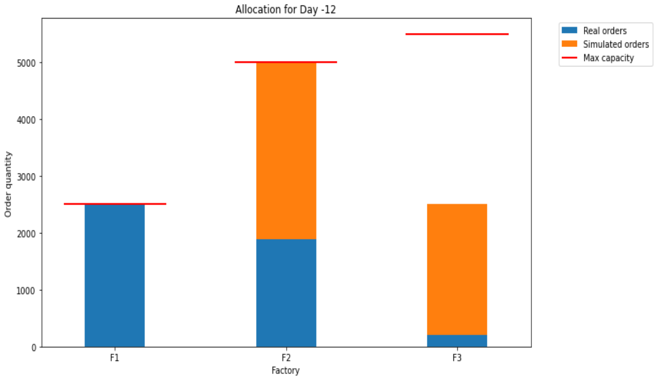


FIGURE 8 Allocation of LD12

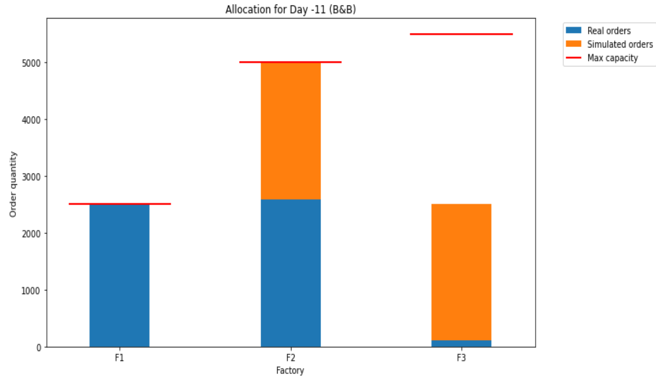


FIGURE 9 B&B allocation of LD11

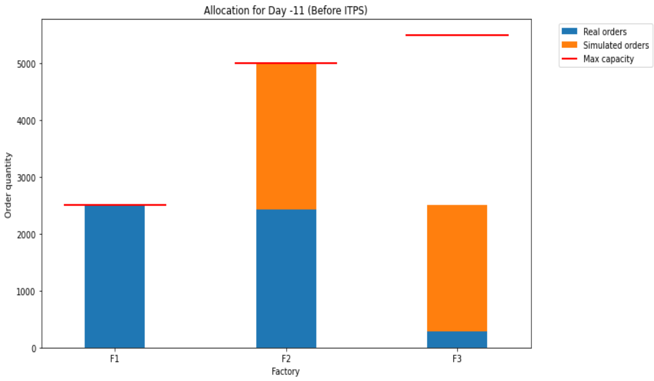


FIGURE 10 Allocation of LD11 before ITPS

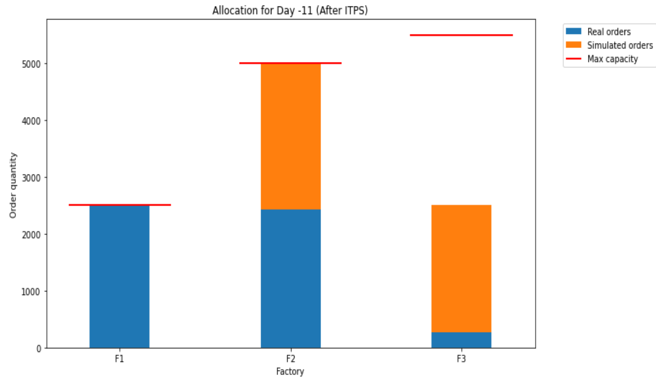


FIGURE 11 Allocation of LD11 after ITPS

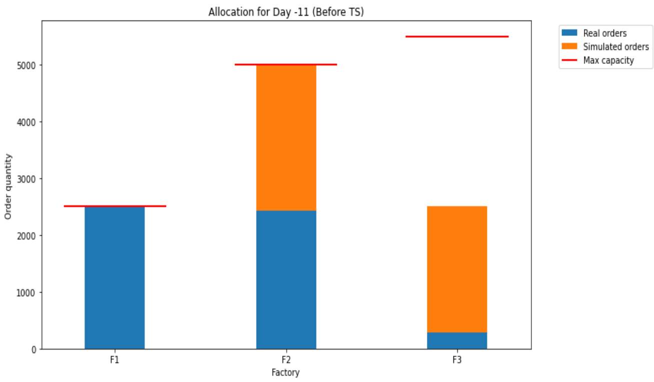


FIGURE 12 Allocation of LD11 before TS

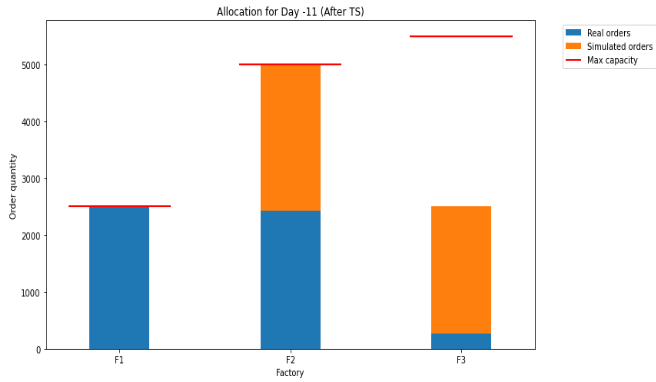


FIGURE 13 Allocation of LD11 after TS

4.2.2 | Scalability test

In this test, each method will be applied to five different order quantities, ranging from 10,000 to 100,000, over two days: LD12 and LD11. Figure 14 compares the optimization times of three algorithms, while Figures 15, 16 and 17 illustrate the gap between WMAPE site and WMAPE global of each method. A result summary is provided in Table 7.

This test confirms B&B is the best method, consistently delivering optimal solutions across all quantities in the shortest time. One noteworthy finding is that when the total order quantity increases, both WMAPEs decrease. This can be explained by the principles of aggregate forecasting which states that forecasts are more accurate for groups of items than for individual items

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

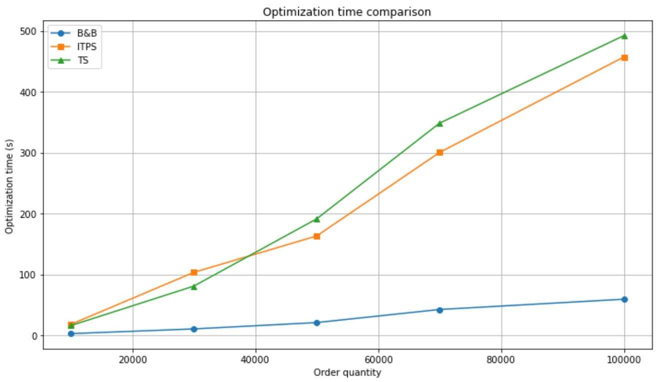


FIGURE 14 Comparison of optimization time

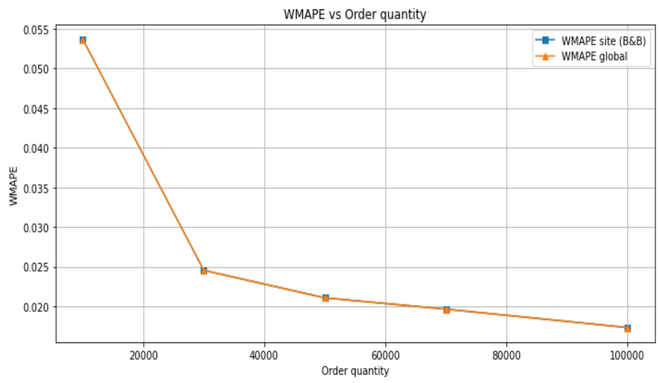


FIGURE 15 WMAPE of B&B

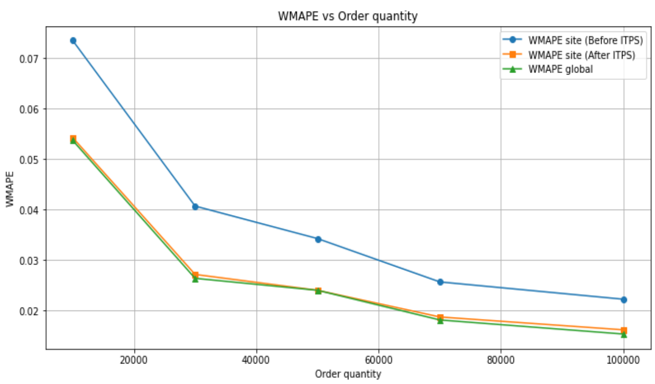


FIGURE 16 WMAPE of ITPS

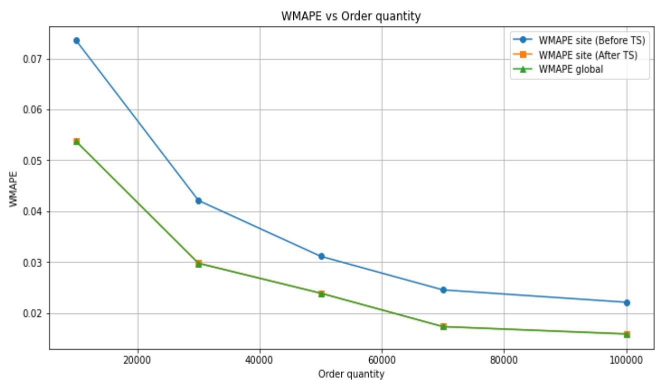


FIGURE 17 WMAPE of TS

because the variability within a group tends to cancel out, leading to more stable and reliable predictions[22]. Moreover, as the order quantity increases, individual fluctuations have less impact on the overall error while in smaller samples, the change of each order has a proportionally larger effect.

TABLE 7 Scalability test results (LD11 and LD12)

Order quantity	Optimization time (s)	WMAPE site (B&B)	WMAPE global
10,000	2.67	0.054	0.054
30,000	10.41	0.025	0.025
50,000	20.84	0.021	0.021
70,000	42.35	0.020	0.020
100,000	59.2	0.017	0.017

Order quantity	Optimization time (s)	WMAPE site (Before ITPS)	WMAPE site (After ITPS)	% Improvement	WMAPE global
10,000	17.91	0.074	0.054	26.21%	0.054
30,000	103.29	0.041	0.027	33.33%	0.026
50,000	163.09	0.034	0.024	29.98%	0.024
70,000	300.28	0.026	0.019	27.14%	0.018
100,000	457.09	0.022	0.016	27.36%	0.015

Order quantity	Optimization time (s)	WMAPE site (Before TS)	WMAPE site (After TS)	% Improvement	WMAPE global
10,000	15.91	0.074	0.054	26.97%	0.054
30,000	80.7	0.042	0.030	29.23%	0.030
50,000	191.19	0.031	0.024	23.24%	0.024
70,000	348.38	0.025	0.017	29.42%	0.017
100,000	492.15	0.022	0.016	28.02%	0.016

4.2.3 | Temporal fixed test

In this test, the best method will be used to address the temporal aspect of BAP. Specifically, B&B will continuously make allocation decisions for a 15-day planning period under ideal conditions, where there are no changes in capacity or real orders. Due to the high computational cost, the order quantity of 10,000 is used. Figures 18 illustrate the increasing proportions of real orders through days. As LD3 approaches, the order composition between consecutive days becomes increasingly similar since each day includes previous real orders, explaining the decreasing trend in error graphs.

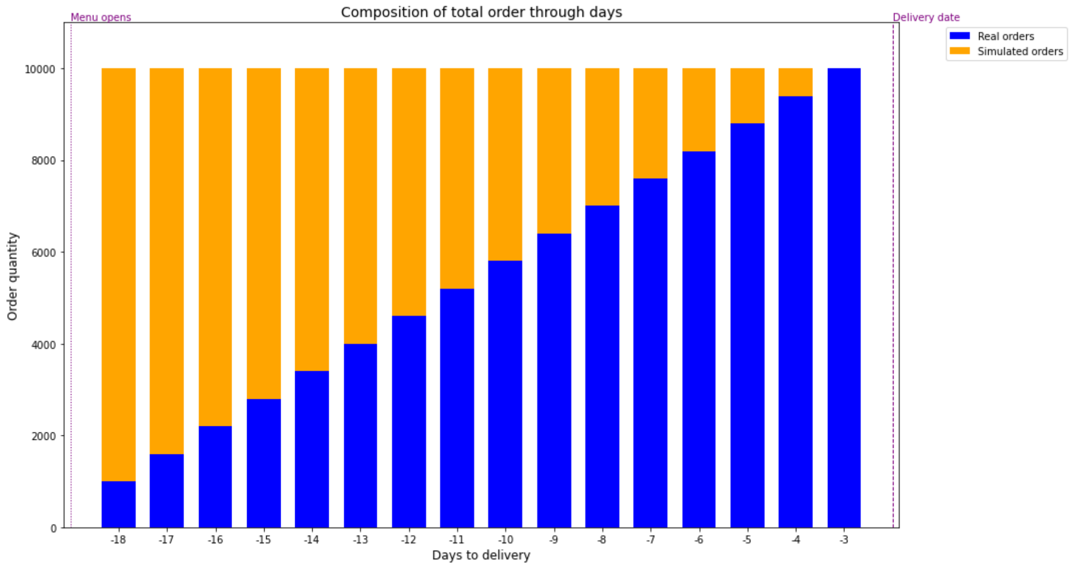


FIGURE 18 Composition of total orders through days

Figure 19 shows WMAPE site when B&B continuously optimizes allocation between two consecutive days. For example, LD16 is optimized based on the B&B allocation of LD17, LD15 is optimized based on the B&B allocation of LD16, and so on. It can be seen that, except for some initial days when the real order proportion is still low, B&B consistently provides optimal solutions from LD13 to LD3. The WMAPE site (Greedy) line shows that without B&B, the recipe differences between two days are significantly higher.

After B&B completes the order allocation for all days, the allocation of each day is compared with the final day's allocation in Figure 20. This comparison evaluates whether day-by-day optimization in Figure 19 can provide 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 recipe differences. The ultimate goal is to minimize the area under WMAPE site curve, ideally making it match the area under WMAPE global curve. Although there is a big gap between the two lines in the early days, it narrows over time. As a result, the area under WMAPE site curve is not much larger than that of WMAPE global curve, confirming the effectiveness of B&B in achieving proxy optimization. Table 8 provides the detailed data of all graphs.

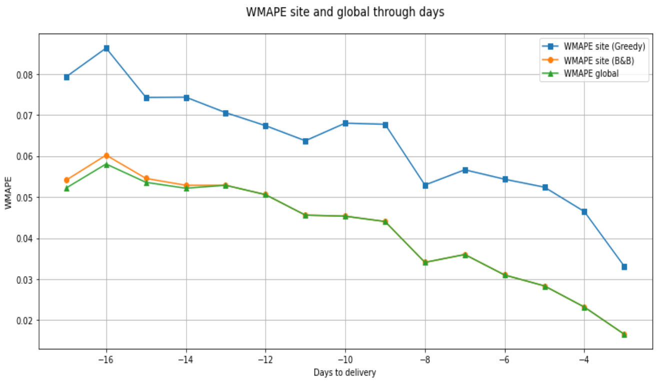


FIGURE 19 WMAPE site and WMAPE global through days

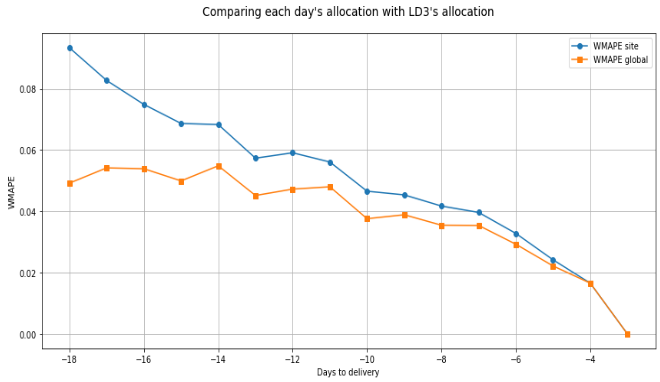


FIGURE 20 Each day's allocation versus LD3's allocation

TABLE 8 Temporal fixed test results

Day	Real orders proportion	WMAPE site (B&B)	WMAPE site (Greedy)	WMAPE global	WMAPE site vs LD3	WMAPE global vs LD3
-18	10%	N/A	N/A	N/A	0.093	0.049
-17	16%	0.054	0.079	0.052	0.083	0.054
-16	22%	0.060	0.086	0.058	0.075	0.054
-15	28%	0.055	0.074	0.054	0.069	0.050
-14	34%	0.053	0.074	0.052	0.068	0.055
-13	40%	0.053	0.071	0.053	0.057	0.045
-12	46%	0.051	0.067	0.051	0.059	0.047
-11	52%	0.046	0.064	0.046	0.056	0.048
-10	58%	0.045	0.068	0.045	0.047	0.038
-9	64%	0.044	0.068	0.044	0.045	0.039
-8	70%	0.034	0.053	0.034	0.042	0.035
-7	76%	0.036	0.057	0.036	0.040	0.035
-6	82%	0.031	0.054	0.031	0.033	0.029
-5	88%	0.028	0.052	0.028	0.024	0.022
-4	94%	0.023	0.046	0.023	0.017	0.017
-3	100%	0.017	0.033	0	0	0

4.2.4 | Temporal variation test

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

1. **Capacity change**: When F1 experiences a significant capacity reduction from 3,000 to 1,000 orders on LD10, a sharp increase in WMAPE site occurs, as shown in Figure 21. This spike is driven by the abrupt reallocation of orders among factories. While B&B demonstrates remarkable agility, swiftly adjusting allocations and achieving the optimal WMAPE site by LD9, the greedy method reveals a notable gap between WMAPE site and WMAPE global after the capacity change, underscoring its inflexibility. The comparison of daily allocations with the final day's allocation in Figure 22 illustrates a smooth convergence toward the optimal WMAPE site value starting from LD10. The sharp reduction in error between LD10 and LD11 signifies a clear transition between two distinct operational scenarios. Consequently, allocation decisions made before LD10 show significant deviations compared to LD3. This abrupt shift highlights B&B's capability to adapt effectively to major changes.

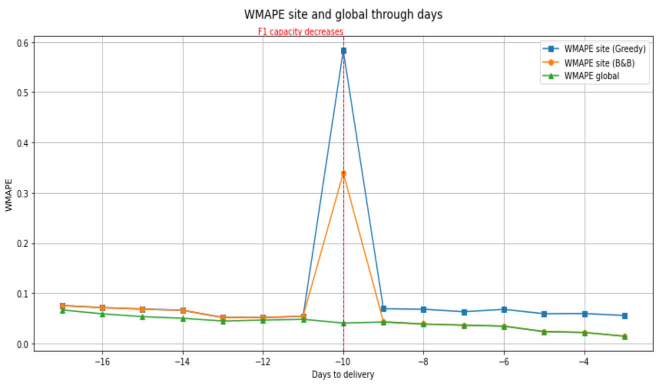


FIGURE 21 WMAPE site and WMAPE global through days

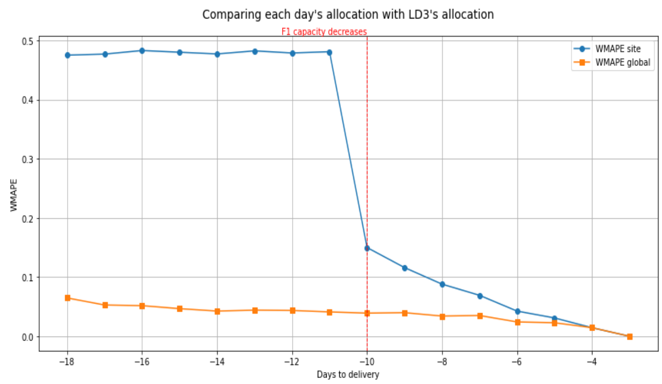


FIGURE 22 Each day's allocation versus LD3's allocation

2. **Order changes**: This study assumes that each day, 5% of real orders are deleted and 30% have recipe adjustments by customers. Figure 23 illustrates how order changes cause fluctuations in both WMAPE site and WMAPE global. Without order changes, two WMAPEs show a clear decreasing trend. However, when changes occur, two WMAPEs fluctuate significantly, depending on the magnitude of these changes. This graph also indicates that B&B can achieve the optimal WMAPE site in any situation.

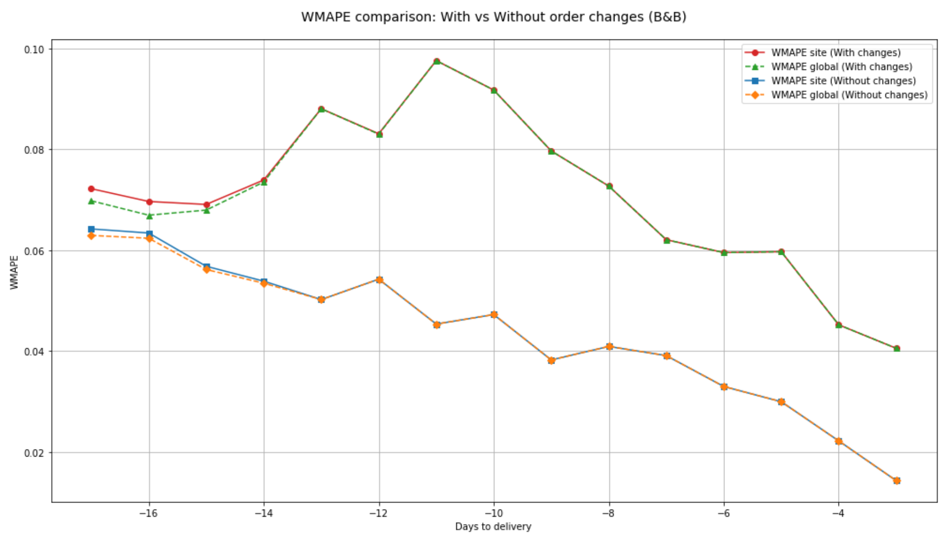


FIGURE 23 WMAPE site and WMAPE global with and without order changes

In this scenario, a straightforward allocation strategy can also be employed since the IDs of real orders remain consistent across days. For instance, if on LD14, order R1 is allocated to F1 and R2 to F2, these same assignments can be retained for LD15. This ID-based approach promotes consistency and minimizes allocation changes for recurring orders. However, there are some problems: recipe changes within each order (for example, R1) may alter its eligibility, potentially making the order no longer suitable for F1. This could confuse the staff at F1 if R1 is still assigned to this factory. Deleted orders can further complicate the situation, as IDs of previous real orders can no longer be retrieved today. Figures 24 and 25 illustrate the above problems by comparing the ID-based method with B&B. The results show that, regardless of order changes, B&B can deliver optimal solutions, matching WMAPE site with WMAPE global. When no order changes exist, the ID-based method can match WMAPE global on some final days when the percentage of real orders is high. However, when there are order changes, the gap between WMAPE site of ID-based method and WMAPE global widens. This underscores the limitations of the simple allocation method and highlights the advantages of more advanced techniques like B&B.

3. **Both capacity and order change**: In the final scenario, we evaluate B&B's performance under simultaneous changes in factory capacities as Figure 26 and real orders. Figure 27 highlights a significant disparity between WMAPE site and WMAPE global achieved by the greedy allocation, whereas B&B consistently delivers the optimal WMAPE site on

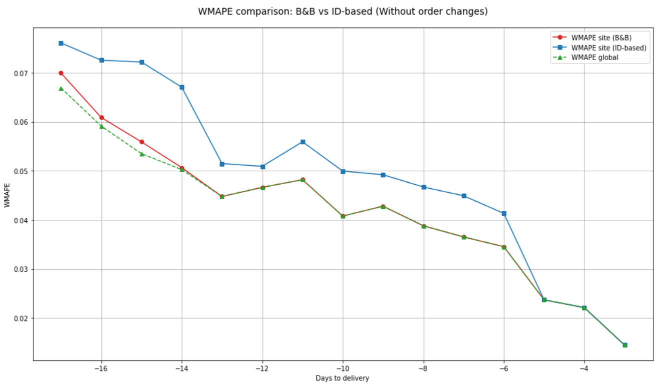


FIGURE 24 B&B and ID-based allocation without order changes

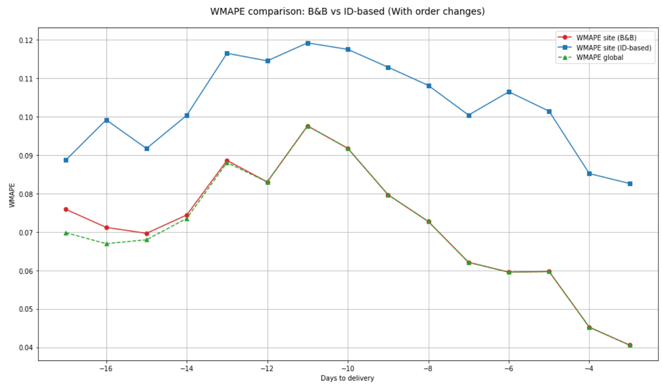


FIGURE 25 B&B and ID-based allocation with order changes

those days without capacity changes. While the greedy method exhibits sharp increases in WMAPE site during capacity adjustments, B&B is impacted only by capacity reductions, which restrict order allocation to the most suitable factories. For example, on LD6, when F2's capacity increases, B&B's WMAPE site remains largely unaffected. This resilience supports a smooth declining trend in WMAPE site from LD10 to LD3 in Figure 28. These findings emphasize B&B's adaptability and reliability in managing scenarios with frequent fluctuations in orders and capacity.

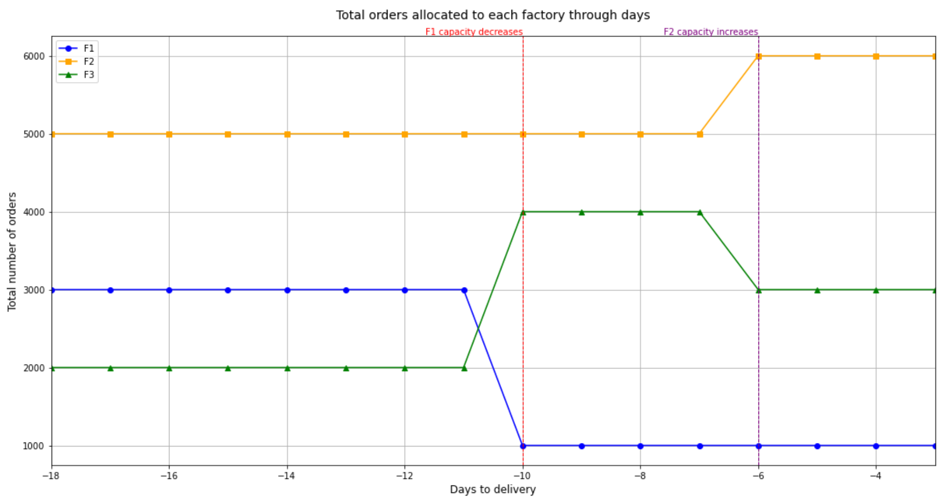


FIGURE 26 F1 and F2's capacity change

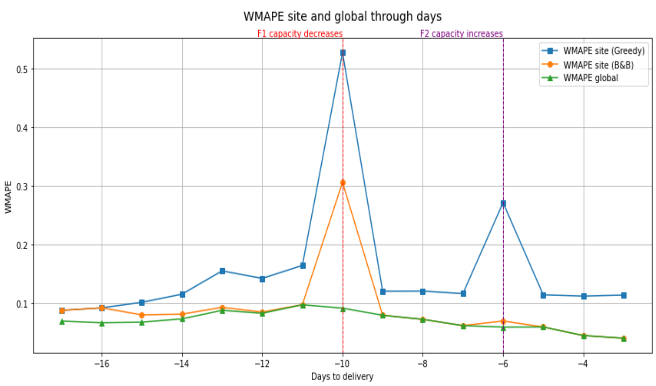


FIGURE 27 WMAPE site and WMAPE global through days

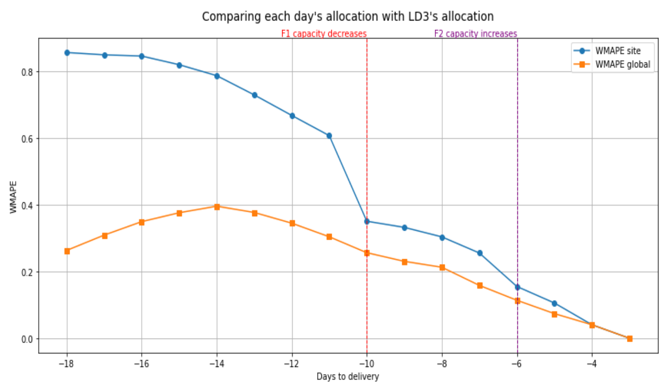


FIGURE 28 Each day's allocation versus LD3's allocation

4.3 | Managerial implications

The results of this study have several implications for decision-makers in meal kit delivery: the superior performance of B&B demonstrates the value of implementing advanced optimization techniques in the box allocation process. Managers should consider integrating this method with their inventory management system to improve operational efficiency and reduce food waste. The variation tests demonstrate the robustness of B&B in handling capacity and order changes, preventing overproduction or underutilization of resources. Managers should prioritize flexible methods to adapt to changes quickly, ensuring minimal impact on overall efficiency. The comparison between B&B and ID-based method reveals the shortcomings of simpler, rule-based approaches. If a similar approach is currently used, managers should be cautious, especially when handling frequent changes. The company's daily generation of new simulated orders introduces unnecessary variability in recipe distribution, potentially increasing errors. A more effective approach would involve preserving well-matched simulated orders from previous allocations, generating new ones only to fill gaps or adjust distribution. This strategy would leverage historical information, leading to more stable allocations and reduced errors over time.

5 | CONCLUSIONS

This study explores the Box Allocation Problem (BAP) encountered by meal kit delivery companies, leveraging insights from BPP and CVRP. The CBC solver demonstrated superior performance by combining primal heuristics like the Feasibility Pump and Coefficient Diving with B&B. Pure heuristic methods, such as Iterated Tabu Search (ITPS) and Tabu Search (TS), also significantly outperformed the greedy approach, with TS nearly matching B&B in delivering optimal results. All tested methods handled up to 100,000 orders within 10 minutes, underscoring their scalability and suitability for large-scale operations.

Extensive experimentation highlighted B&B's strength in quickly minimizing day-to-day recipe variations, ensuring a smooth transition from soft to hard allocations. Its robustness was evident in adapting to fluctuating factory capacities and real orders. Scalability tests revealed that larger order volumes significantly reduced errors, indicating that scaling operations or consolidating more orders into a single allocation cycle could further improve accuracy.

However, this research has some limitations. The current model focuses on daily optimization and lacks a dynamic, long-term planning perspective. Future research could integrate horizon-wide strategies to improve overall efficiency. Moreover, metaheuristics like Genetic Algorithms or Large Neighborhood Search remain unexplored and offer potential for further performance gains. Validating the model with real-world data and designing algorithms to ensure consistency in simulated orders across runs could also reduce errors and advance optimization for meal kit delivery and related applications.

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. Clear Spider (2024) *How Meal Kit Delivery Services Manage Their Supply Chains*. Available at: https://clearspider.net/blog/meal-kit-delivery-supply-chains/.

4. 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.

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 *Emerging theory, methods, and applications*. INFORMS, 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.

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/2020/06/23/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.

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. 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.

17. Loic, G. (2024) Internal material.

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/10.49550/arXiv.2301.09943.

21. Radcliffe-Brown, A.R. (1939) *Taboo*. Cambridge: Cambridge University Press. Available at: https://www.amazon.co.uk/Taboo-Frazer-Lecture-R-Radcliffe-Brown/dp/1107695791.

22. Reid, R. Dan and Sanders, Nada R. (2012) *Operations Management: An Integrated Approach, 5th Edition*. Available at: https://www.oreilly.com/library/view/operations-management-an/9781118122679/ch8-sec004.html.

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. (2002) *The Vehicle Routing Problem*, SIAM Monographs on Discrete Mathematics and Applications. Available at: https://api.semanticscholar.org/CorpusID:209742615.

26. Woop (2024) *Home*. Available at: https://woop.co.nz/.

27. Zangl, G., Graf, T. and Al-Kinani, A. (2006) ‘Proxy Modeling in Production Optimization’, *Paper presented at the SPE Europec/EAGE Annual Conference and Exhibition*, Vienna, Austria, 12-15 June. Available at: https://doi.org/10.2118/100131-MS.

How to cite this article: Thu T. M. N., Alain Z., and Loïc G. Box allocation optimization in meal kit delivery. *Journal.* 2024;00(00):1–18. https://doi.org/...