

Improved Packing Strategy for Distribution Centers to Reduce Freight Cost

by

Bowen Zeng

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of
Master of Engineering in Advanced Manufacturing and Design

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2019

©Bowen Zeng, MMXIX. All rights reserved.

Author
Department of Mechanical Engineering
August 15, 2019

Certified by
Stephen C. Graves
Abraham J. Siegel Professor of Management
Thesis Supervisor

Accepted by
Nicolas Hadjiconstantinou
Chairman, Department Committee on Graduate Theses

This page was intentionally left blank.

Improved Packing Strategy for Distribution Centers to Reduce Freight Cost

by

Bowen Zeng

Submitted to the Department of Mechanical Engineering
on August 15, 2019, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Advanced Manufacturing and Design

Abstract

In this thesis, we designed and implemented a data-driven packing strategy for distribution center outbound packing activities to reduce freight cost and carbon footprint. The strategy consists of two parts. First, I proposed an Carton Combination method, an algorithm that can select any predetermined number of distinct cartons from a large carton pools (over 1000 options) to be used for outbound shipment packing such that the annual total wasted air content inside the shipment is minimized. Second, I proposed an Carton Selection algorithm, which can determine the best carton, from the carton options chosen by the Carton Combination method, for an incoming order with known dimensions. The entire packing strategy prototype was implemented by MATLAB R2019a; the prototype was tested with the 2018 outbound shipment data from Waters Corporation Global Distribution Center (GDC) and the simulation showed that the annual averaged shipment air percentage was reduced from 60% to 40%, which projects to an annual operation cost saving of 83,000 USD and carbon dioxide emission reduction of 20 *ton*. The data-driven packing strategy has a potential to be scaled up and implemented via an industrial environment such as SAP ABAP.

Thesis Supervisor: Stephen C. Graves
Title: Abraham J. Siegel Professor of Management

This page was intentionally left blank.

Acknowledgments

I would like to extend my sincere and heartfelt obligation towards all personages who have helped me in this endeavor. Without their active guidance, help, cooperation and encouragement, I would not have made headway in the thesis. This thesis would not have been possible without:

- Li-Chung Pan, Logistics Project Manager at Waters Corporation, for his constant guidance, inexhaustible patience, and excellent management support.
- My colleagues, Dehui Yu and Jessica Harsono, for our great collaboration, discussion, and mutual support and encouragement.
- Our thesis advisor, Prof. Stephen Graves, for his guidance on sub-project development, algorithm skeleton establishment, and long-term vision projection.
- Dan Welch, VP of Global Supply Chain, and Kathleen Wright, GDC Manager at Waters, for their constructive feedback and insights on our work on along the way.
- Gabriel Kelly, Purchasing Manager at Waters, for his great support on MIT-Waters Collaboration initiation, and the "corporation alignment" lesson.
- Other team members in Waters Corporation from GDC, EDC and Global Manufacturing, for their great support and advice that made every visit to Waters worthwhile.
- My MEng peers, Diarny Fernandes, Siyang Liu, Paul Zhengyang Zhang, Stratos Moskofidis, and Steve Ratner, without whom I would lose a lot of great memories of 2018-19.
- Last but not least, my family, who always stands behind me 120%.

This page was intentionally left blank.

Contents

1	Introduction	13
1.1	Problem Statement	13
1.2	Motivation	14
1.3	Objective	14
1.4	Scope	14
1.5	MIT Team and Work Distribution	15
2	Background	17
2.1	Waters Corporation	17
2.2	Waters/MIT Collaboration and Past Projects	18
2.3	Waters Global Distribution Center	19
2.3.1	Warehouse Functions	20
2.3.2	Warehouse Layout	20
2.3.3	Worker Roles	23
2.3.4	SAP and Warehouse Management System	25
2.3.5	GDC Outbound Operations Value Stream Mapping	25
3	GDC Packing Problem	29
3.1	Waters Corporation Goals	29
3.2	Current Packing Operations	30
3.3	Shipping Cost Calculation	32
4	Baseline Establishment	35

4.1	Baseline Data	36
4.2	Baseline Assumptions	37
4.3	Baseline Analysis Results and Opportunities	38
5	Carton Combination Method	39
5.1	Problem Simplification	39
5.2	Ideal Solver	40
5.2.1	SKU Dimensional Constraint	44
5.2.2	Data Input	47
5.2.3	Ideal Solver Results and Challenges	50
5.3	Quick Solver	51
5.3.1	Box Pool Down-selection	51
5.3.2	Constraint Relaxation	51
5.3.3	Carton Worth Down-selection Method	53
5.3.4	Results and Discussion	55
6	Carton Selection Algorithm	59
6.1	Algorithm Overview	59
6.2	Updated Packing SOP	61
7	Conclusion and Future Works	65
A	MATLAB Script	69

List of Figures

2-1	Waters' GDC layout.	21
2-2	Waters' outbound operations value stream mapping.	26
3-1	U.S. express package (public) rates example of FedEx.	33
4-1	Data structure for the GDC packing baseline establishment.	36
5-1	Screen-shot of box info from Packaging Supplies.com.	48
5-2	Order quantities assigned to each carton in pool.	52
5-3	Order quantity assigned to each carton in the pool with the k -means clustering and carton combination results called out.	54
5-4	Averaged shipping air percentage vs. size of carton set.	57
6-1	Carton selection algorithm flow diagram.	60
6-2	New GDC packing SOP.	62

This page was intentionally left blank.

List of Tables

3.1	Current carton SKUs in GDC.	32
5.1	Example of master order info database.	48
5.2	Optimized carton options via ideal solver.	50
5.3	Carton SKU combination via quick solver—set of 12 cartons.	55
5.4	Carton options via quick solver—set of 20 cartons.	56

This page was intentionally left blank.

Chapter 1

Introduction

1.1 Problem Statement

Improving the efficiency within a distribution center by reducing waste and optimizing process flow can result in significant cost savings regarding distribution center operations. One area with a large potential for improvement is outbound order packaging. Shipping costs are calculated by courier service providers based on the actual package weight or the dimensional weight, whichever is higher. With a limited number of box sizes stocked and available in the distribution center to package orders, some orders will inevitably be shipped in a box that is larger than needed to accommodate the products inside. This unnecessary extra volume means that some portion of the cost is due to shipping the air within the box—a cost that can be eliminated or reduced if a better packaging choice were to be utilized.

The team explored and proposed three potential solutions to reduce the empty volume within shipping packages and thus reduce freight costs. The three solutions are: introducing new packaging options such as envelopes and paks, creating a data-driven box selection strategy/algorithm to assist workers in choosing the most economical stock box for packaging, and creating an integration plan for adding custom box-making technology to the current packing process flow.

1.2 Motivation

The motivation of this project is to reduce the waste caused by the unnecessary empty volume within shipping packages. This packing problem is quite common for distribution centers and there is huge waste in freight cost, as well as excess corrugate usage and more CO₂ emission. Therefore, it is valuable to develop some potential solutions, followed by detailed evaluation to quantify the impacts and difficulty of implementation.

1.3 Objective

The key objective for this thesis is to evaluate the three proposed solutions for packing problem from different perspectives to provide recommendations for distribution centers. An evaluation method for these outbound improvement projects at distribution centers will be established, and it will contain three main components—financial, implementation and sustainability. Then this method will be utilized to evaluate the three solutions proposed for Waters’ Global Distribution Center (GDC) specifically as a sample case study.

1.4 Scope

The scope of this thesis is limited to the proposal and evaluation of three solutions for domestic shipping at Waters Global Distribution Center. Special shipments, such as hazardous material and cold chain shipments, as well as international shipments (account for less than 10% of GDC total shipments) were not considered. Furthermore, this thesis only analyzed three potential solutions from three main perspectives. There are other factors that may need to be considered for actual execution.

1.5 MIT Team and Work Distribution

The MIT team consisted of 3 people: Jessica Harsono, Dehui Yu, and Bowen Zeng, the author of this thesis. The team analyzed the problem, established the baseline, and proposed three potential solutions together. Thus Chapter 1, 2 and 3 of this thesis were co-authored by the whole team. The author of this thesis focused on the development of the carton combination/selection algorithm and its implementation. Dehui Yu's thesis focuses on the adoption analysis of envelops/paks and the evaluation of the solution. Jessica Harsonos thesis covers the operational analysis regarding the implementation of the custom box maker.

This page was intentionally left blank.

Chapter 2

Background

The Waters Corporation is the sponsor and host company for this thesis project and is the subject of the research conducted. This section provides background on the Waters Corporation as a business and breaks down the functions within their Global Distribution Center (GDC). The history of MITs collaboration with Waters and relevant past thesis projects are briefly discussed to provide context and scope for this thesis.

2.1 Waters Corporation

Waters Corporation is the world's leading specialty measurement company focused on improving human health and well being through the application of advanced analytical science technologies. Founded by Jim Waters in 1958, Waters serves life sciences, food sciences, and materials sciences through a connected portfolio of chromatography, mass spectrometry, and thermal analysis innovations. With approximately 7,200 employees worldwide, Waters operates directly in 31 countries, including 15 manufacturing facilities, 3 distribution centers, with products available in more than 100 countries. The company's main headquarters is located in Milford, Massachusetts, USA. Waters products include high-performance liquid chromatography (HPLC), ultra-performance liquid chromatography (UPLC) and mass spectrometry (MS) technology systems and support products. In addition, the company also provides thermal analy-

sis, rheometry and calorimetry instruments, as well as other software-based products.

Waters is a publicly-traded company listed on the New York Stock Exchange. The market cap was 13.99 billion USD on August 2nd, 2019. For the 2018 financial year, Waters had 2.4 billion USD revenue and 599 million USD free cash flow. Of the 2018 sales, 38% came from Asia (18% from China and 7% from Japan), 35% from America (28% from the US), and 27% from Europe. From the end market perspective, 56% of the revenue was from the pharmaceutical industry, 19% from material sciences and 15% from the food and environment market.

2.2 Waters/MIT Collaboration and Past Projects

Waters and the MIT Master of Engineering in Advanced Manufacturing and Design program have collaborated annually since 2013 on projects ranging from product design, to research and development for new processes, to manufacturing process control, to operations improvement, and to supply chain. Every academic year, a team of MIT student researchers from the program address projects from various departments at Waters to realize current strategic goals and provide new perspectives for the future of the company. Past projects have resulted in patent-worthy scientific applications and have helped bring fresh innovative ideas into the history of scientific development existing at Waters. Having an outsiders opinion helps them to gain fresh viewpoints on their challenges. In addition to the student team, a faculty advisor offers subject matter expertise to guide students' work. At the same time, the students develop contextual thinking by getting a holistic view of the business and gaining tangible industry experience.

Although the Waters/MIT collaboration has existed since 2013, this thesis marks the second year that projects have been carried out at Waters new Global Distribution Center located at Franklin, MA. This new warehouse is larger and can support more capacity than the previous warehouse located at Waters headquarters in Milford, MA. In 2018, the previous MIT student team led an exploration phase characterizing the new GDC and proposing various improvement pathways for project categories

including optimization, digitization, automation, and standardization. The two main projects they ultimately pursued were radio frequency identification (RFID) for products within GDC and the creation of a heat map suggesting how to allocate SKUs storage locations within GDC in order to strategically ensure that the fastest moving products were the most easily accessible to workers [8]. These projects focused primarily on the picking side of warehouse operations, while this year's projects focus more on the operations associated with packing orders. The definitions of picking and packing are elaborated later in this chapter.

The projects chosen by the MIT students during any particular year typically align with the state of the company and in areas where immediate short-term to mid-term returns are desired. This keeps the projects relevant to students' academic field and ensures that the necessary resources are readily available to realistically carry out tasks and provide the most value to the company.

2.3 Waters Global Distribution Center

The Global Distribution Center (GDC), where this thesis was conducted, is one of Waters' three distribution centers. The corporation operates two other distribution centers internationally: the European Distribution Center (EDC) in the Netherlands and the Asian Distribution Center (ADC) in Singapore. GDC is the largest among them and is the only distribution center in the United States. In addition, GDC functions more comprehensively compared with the other two as it ships both to the other two distribution centers as well as to customers globally. GDC is a 15-minute driving distance from the global headquarters of Waters in Milford, MA, where the original GDC was located. With this relocation in October 2017, GDC was scaled up to a 56,000-square-foot modern facility with various types of storage to match the pace of the company's growth. As of August 2019, GDC maintains over 14,000 distinct SKUs in over 20,000 separate storage locations consisting of 97% of the warehouses designed capacity (based on the current layout). The corporation projects a compound annual growth of 6% in terms of outbound flows in the next

five years, and such a fast growth rate has already started to impose both storage capacity and operation challenges to GDC.

2.3.1 Warehouse Functions

GDC functions as a warehouse to store all of Waters current inventory within the United States. SKUs are received from vendors and production locations and are used to re-stock the active SKUs in their current storage locations or are placed in any available space if the designated locations are full. Occasionally, new products will be procured and must also be stored. GDC must store enough safety inventory to provide a high level of service to customers and keep the order lead time low. All of the storage locations are labeled using a chronological numbering system that spans the entire length of the warehouse and lets pickers know the location of an item on their order list.

GDC also serves to distribute products to customers and fulfill orders both domestically and internationally. Customers are typically laboratories, engineers, scientists, and smaller distributors. Products in GDC are also distributed to the other two distribution centers (ADC and EDC) depending on the inventory levels and current needs. Domestic orders are typically smaller and more common, with only a small portion of freight orders¹, while international orders are almost always large freight orders. Orders are almost always picked, packed and shipped on the same day that the orders are placed. Orders are shipped via a variety of methods such as air or ground with multiple carriers such as FedEx and UPS, depending on the request of the customer.

2.3.2 Warehouse Layout

The GDC has multiple functional areas, and its layout is shown in the following figure.

¹Heavy orders shipped via pallets and wrapped with stretch wrap film

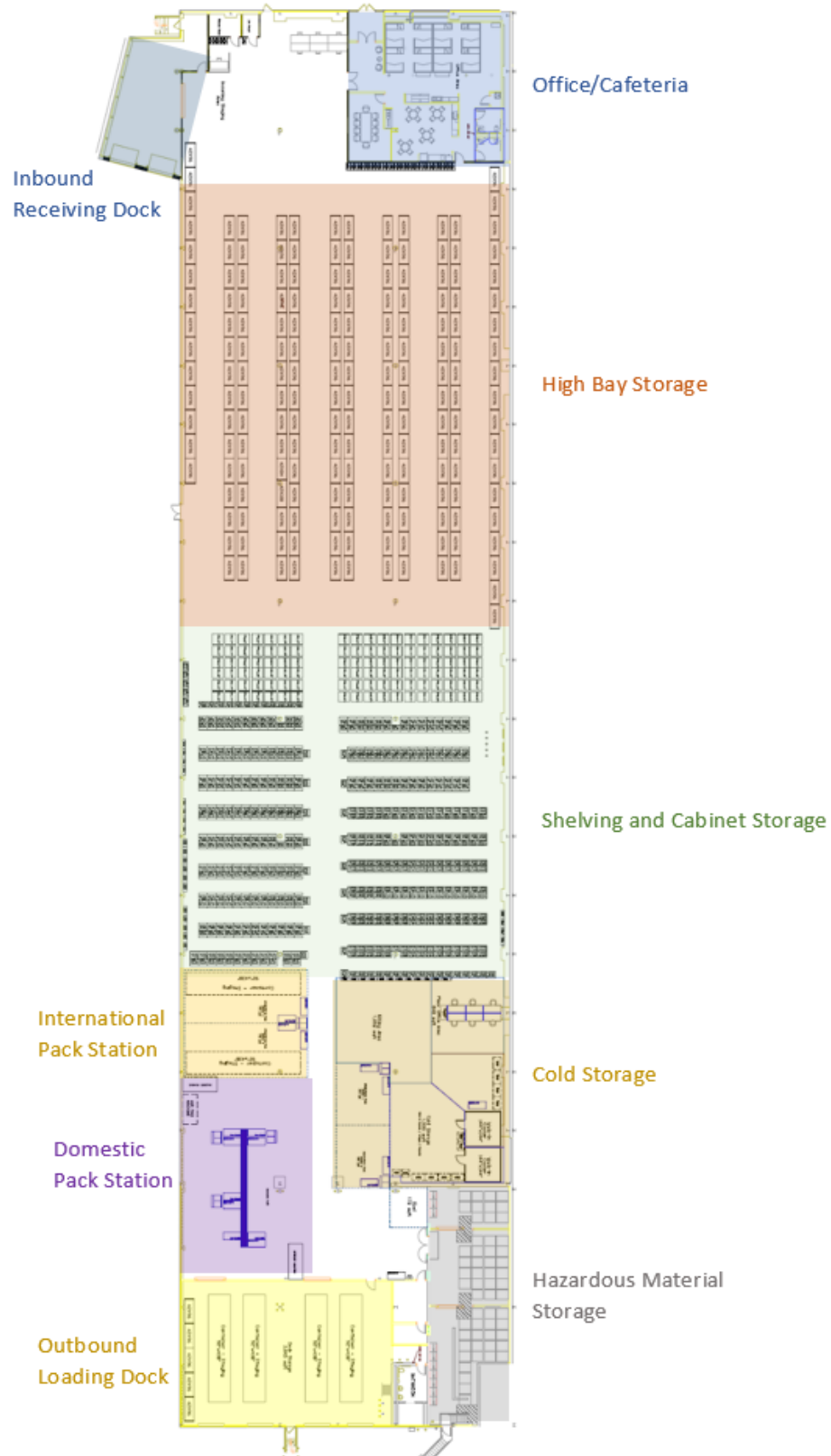


Figure 2-1: Waters' GDC layout.

The detailed description of each functional area is listed below:

- *Office/Cafeteria:* The GDC manager, leads, supervisors, and other employees sit at the front of the building. There are cubicles for computer and desk work and a conference room to hold meetings.
- *Inbound Receiving Dock:* All arriving products enter GDC through the inbound receiving dock where they are unloaded off trucks. The items are transported on pallets or in carts and are scanned and digitally documented by receivers before being placed within GDC.
- *High Bay Storage:* High bay storage is used for large items such as instruments, machines, and pallets of products. They span multiple stories high and require a forklift to access some of the topmost SKUs. The high bay area is the storage space that is furthest from the pack stations and requires the most walking to get to by the pickers.
- *Shelving and Cabinet Storage:* The lower shelving and cabinet storage is used for smaller SKUs such as columns, spare parts, and literature/software. These locations do not require a forklift to access and are closer to the pack stations.
- *Cold Storage:* The cold storage area holds all of the products that require freezing or refrigeration. It also holds the necessary equipment needed to properly package these products for shipping.
- *Hazardous Materials Storage:* The hazardous materials storage area holds items that require special caution, such as chemical or flammable substances. Most of these products also have an expiration date and must be monitored for quality. The area can be isolated from the rest of the distribution center in case of fire or flood.
- *International Pack Station:* The international pack station deals with all orders being shipped outside the United States. There is more floor space to accom-

modate the large shipping containers and pallets used for freight orders and no conveyor belt since there are fewer shipments.

- *Domestic Pack Station:* The domestic pack station sees greater package throughput and takes care of all orders shipped within the United States. These packages are smaller and can fit on a gravity conveyor, which is shared between the pack benches. Orders are packaged at the benches and sent down the conveyor for shipping.
- *Outbound Loading Dock:* Once orders are packaged and ready for shipment, they are sent to the outbound loading dock either by cart or by forklift/pallet jack. Trucks of multiple delivery carriers come at specified times of the day to pick up packages for shipment.

2.3.3 Worker Roles

The GDC employs around 20-30 full-time workers at any given time, with some temporary employees hired during busy seasons or for special projects. Workers at GDC have various roles as following:

- *Distribution Operations Supervisor:* Responsible for managing the operations of GDC and making sure all the procedures are implemented effectively.
- *Inventory Control Specialist:* Responsible for conducting daily reporting and making inventory adjustments and removing any product from the warehouse that has been identified as reaching the minimum remaining shelf life.
- *Section Lead:* Responsible for supervising all processes within the specific section and providing necessary instructions/support for material handlers, such as order dropping and work assignment.
- *Material handler:* Responsible for the actual material handling processes ranging from receiving, picking, packaging to shipping. The workers are cross-trained

but a material handler will only perform one duty based on the work assignment within a specific time period to avoid mistake propagation. The specific handling activities are as follows:

1. Receiving—Worker unloads incoming pallets of product from trucks into GDC via the receiving dock and scans items into SAP to verify their delivery. They take the product and store it in available locations within GDC.
2. Picking—Worker receives printed copies of orders from the section lead detailing which SKUs to retrieve from GDC storage. They walk with a cart and can fulfill multiple orders for every trip to and from the packing stations. Occasionally, certified forklift drivers are necessary to retrieve SKUs stored in the high bay area.
3. Packing—Worker stands at a packing bench stocked with all of the necessary materials to pack an order. They receive a cart full of picked orders and begin processing them one by one. They scan/verify each item of the order into SAP, select and assemble a stock sized cardboard box, place the order into the box, add plastic pillows as void fill, seal the box with packing tape and print and stick the delivery label for the order onto the box. They put the completed box on a gravity conveyor that then goes to the shipping station.
4. Shipping—Worker stands at a bench with a weight scale and computer with a special shipping program located at the end of the packing area. They take all of the packaged orders and scan the delivery label, weigh the order, and print out the appropriate shipping label. If there are missing fields in an orders shipping information or any incorrect information, then the worker has to correct it before completing the shipping label for that order. Once the shipping label is placed on the package, it is fully ready to be sent to the outbound loading dock where it will be picked up by truck.

2.3.4 SAP and Warehouse Management System

SAP, which stands for systems, applications, and products in data process, is a German multinational software corporation that makes enterprise software to manage business operations and customer relations. Its software shares the same name with the company. GDC widely uses SAP as a warehouse management system (WMS) to link the inbound and outbound activities:

1. Material receiving (inbound)
2. Order creation (outbound)
3. Order verification (outbound, packing station)
4. Internal shipping label generation (outbound, packing station)
5. Shipping label generation (outbound, shipping station)
6. Customer relation report generation and analytics (enterprise level)

SAP links data among all above activities to ensure each SKU and each order are traceable during operation. Using SAP as a WMS is to create a digital twin of all storage locations. Because of the versatility of SAP, it often integrates with other software and modules to fulfill special functions. For Waters GDC, SAP integrates with CMS WorldLink to centrally manage domestic and international shipments by multiple carriers.

2.3.5 GDC Outbound Operations Value Stream Mapping

Value stream mapping, also called material and information flow mapping at Toyota, is a method of lean management that is used to analyze the current state of a process and then to design an improved future state to reduce waste. A value stream focuses on areas that add value to a product or service [9].

In order to have a better understanding of GDCs current practices, a value-stream map was made by the team as shown in the following figure. Since the scope of

this project is within the domestic outbound operations of GDC, the manufacturing and inbound processes were not included. The components of the GDC outbound operations value stream mapping are shown in the figure below and presented as follows:

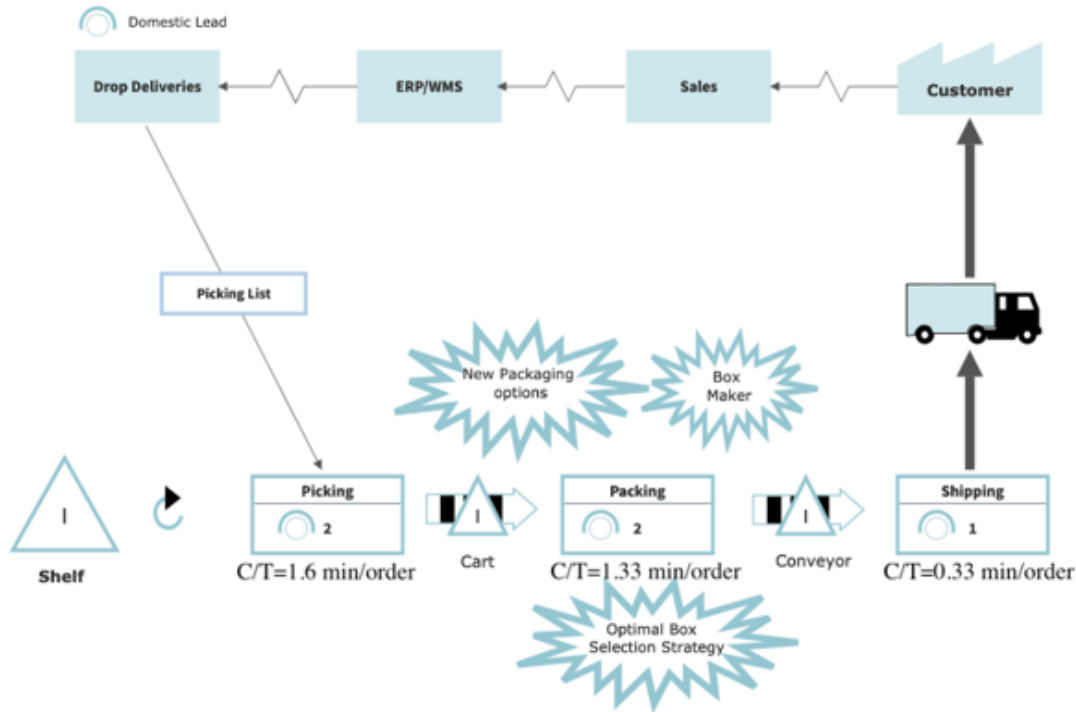


Figure 2-2: Waters' outbound operations value stream mapping.

- *Information Flow:* After customers (either external or internal) place orders, the order information will be transferred to Waters sales department, and then to the Enterprise Resource Planning (ERP) system and WMS hosted by SAP. The domestic lead at GDC then will drop the deliveries from the server and print out hardcopy picking lists for the orders.
- *Material Flow:* The picker will grab one cart and pick the corresponding products from different storage locations based on the assigned picking lists from the domestic lead. After all the items are picked, the picker will push the cart to a packing station. The packer will then verify the products via WMS and pack them accordingly. The packed boxes will be put on the conveyor belt and

moved to the shipping station. A material handler will weigh the packed boxes and print out the actual shipping label and sort them into different bins for couriers to collect.

The above figure also shows the cycle time (C/T) for each step. The cycle time was measured per person instead of per station, because the number of workers at each station changes. The average throughput rate under the current operation is 55 orders/hour.

- *Personnel:* For normal operation, on average there are two pickers, and two packers working together for domestic deliveries, and there is one worker that handles the shipping process.

This page was intentionally left blank.

Chapter 3

GDC Packing Problem

Improving the efficiency within a distribution center by reducing waste and optimizing process flow can result in significant operating cost savings. One area with a large potential for improvement is outbound order packaging. The team explored and developed potential solutions to reduce the box volume per order and in turn reduce shipping costs. This involved introducing new packaging options such as padded envelopes, creating algorithms to assist packers in choosing the most economical stock box while packaging order and creating an integration plan for adding custom box-making technology to the current packing process flow.

3.1 Waters Corporation Goals

We describe here three company goals as they relate to the operation of the GDC.

- *Reduce Freight Cost:* Waters spends more than 10 million USD on logistics every year. Therefore, even a small percentage of improvement will yield a large amount of savings. With a finite number of box sizes stocked and available in the distribution center to package orders, some orders will inevitably be shipped in a box that is larger than needed to accommodate the SKUs inside. This unnecessary extra volume means that some of the cost is used to pay for shipping the air within the box, a cost that can be eliminated or reduced if orders were able to be packaged in smaller boxes.

- *Increase GDC Capacity and Throughput:* GDC is currently operating at 97% capacity in terms of storage space, and the rental contract for the facility will not terminate until 2027. There is limited space to conduct any further expansion work. Therefore, Waters is keen on finding methods to accommodate future growth, which is projected to be 6% annually, by more efficiently utilizing GDC resources. In addition, during every EOQ, there is a sharp demand surge, a situation that requires GDC to hire additional contract operators to match throughput. Hence, Waters also aims to improve the efficiency of GDC operations and internal logistics. Typically, if fast-moving products that are commonly ordered by customers can be placed closer to the pack stations, then the pickers do not have to walk as far to retrieve them. The heat map project from the previous year optimized the locations of the SKUs to minimize the walking time of pickers. However, the labor hours required to re-arrange the entire warehouse have not yet been put in and the proposed layout has not yet been implemented.
- *Sustainability:* Sustainability, defined by the United Nations as development that meets the needs of the present without compromising the ability of future generations to meet their own needs is also an emphasis of Waters Corporation. Waters sustainability efforts focus on six areas: customer collaboration, innovative solutions, value chain management, environment and safety responsibility, community engagement and employment. For this project, the applicable sustainability impact will be mainly on the environment, and the specific target is to reduce box sizes for packing orders. Smaller boxes yields two sustainability benefits. One is to reduce corrugated cardboard usage; the second is to reduce the sipping volume, which can then decrease CO₂ emission.

3.2 Current Packing Operations

At GDC, finished goods are packaged in a manner that conforms to customer requirements and corporate goals: minimizing possible transport damage, being cost-

effective, providing accurate processing of all customer orders, while maintaining efficient inventory control.

GDC ships both domestic and international orders, with international shipments accounting for less than 10% of GDC outbound shipments. International orders are typically pooled by location and sent out in large batches using “D” and “E” size containers¹. The packers already have an efficient way to trim the top of these containers and minimize their total volume. More opportunity lies in the domestic shipments, due to the greater number of orders and the room for process improvement. Therefore, the scope of this thesis was narrowed down to domestic outbound operations.

The Standard Operating Procedure (SOP) for domestic order packing is as follows: Finished goods will be overpacked by selecting a shipping box that allows enough space on all sides. Packing material (air pads) is then added to fill us voids and to protect the product during shipping. Given the available box sizes, every effort is made to minimize the overall size of the package, which minimizes freight costs and the use of shipping supplies. The primary objective of ensuring adequate protection of the finished goods will not, however, be compromised. Larger products that are pre-packaged in suitable shipping boxes, usually instruments and products that will not fit inside one of the standard shipping boxes, do not need to be over packed. Certain products are marked on the product label as “Direct Shipment Overpack Not Required”. These products can be shipped as packaged [6].

Currently, GDC uses 15 standard-size stock boxes, and the detailed names and dimensions are shown in the following table. For domestic and Canadian shipments, only the top 12 box selections are in use, while the FED Box, D-container and E-container are used for international shipments only. In addition, the FedEx Small, Medium and Large boxes can only be used for FedEx express service.

¹For container details, please refer: <https://www.airseacontainers.com/d-container-gaylord-box-with-pallet.html#page=page-1>.

Box Types	Dimensions (<i>in</i>)		
	L	W	H
J12	11	8	5
J14	13	10	6
J18	18	14	8
J22	20	16	10
J64W	26	16	14
Square	13	13	7
Small Column	16	6	6
Large Column	17	9	9
MD262020	26	20	20
Small FedEx (FedEx Express Only)	13	11	2
Medium FedEx (FedEx Express Only)	14	12	3
Large FedEx (FedEx Express Only)	18	13	4
Fed Box (Not for Small Pack)	40	20	26
D Container (Not for Small Pack)	50	43	61
E Container (Not for Small Pack)	47	32	31

Table 3.1: Current carton SKUs in GDC.

As evidenced by the SOP, Waters does aim to find the optimal size box for each delivery from this set of boxes. However, the box selection process is purely based on the packer’s experience and personal judgment. Therefore, the team proposes multi-dimensional solutions on hardware, software and management perspectives to systematically improve the packing operations.

3.3 Shipping Cost Calculation

The shipment charging method for GDC is summarized in this section. Waters has many courier service providers, including FedEx, DHL, UPS, Horizon Air Services and etc. Since FedEx shipped about 90% of GDC outbound shipments in 2018, this shipping cost explanation is based on FedEx’s method, while other companies’ methods are similar.

FedEx rates are mainly based on the shipping zones. Shipping zones are categorized based on the distance between the sender address and recipient address. For example, for shipments within the contiguous United States, 0-150 miles are catego-

alized as Zone 2, 151-300 miles as Zone 3, 301-600 miles as Zone 4, and so on². The following figure is an example of the FedEx published shipping rates for Zone 2, and all other zones' shipping rate calculation follows the same format. As following figure shows, for each zone, the freight rates are decided by the service type, packaging, and billable weight. The billable weight is either the actual weight or the dimensional weight of the package, whichever is larger. The dimensional weight is calculated as the package volume divided by a specific dimensional factor defined by the annual contract between Waters and FedEx.

U.S. Express Package Rates: Zone 2¹

Shipments moving generally 0–150 miles from origin to destination anywhere in the contiguous U.S.

Delivery Commitment ²	Next day by 8 or 8:30 a.m.	Next day by 10:30 a.m.	Next day by 3 p.m. ³	2nd day by 10:30 a.m.	2nd day by 4:30 p.m. ³	3rd day by 4:30 p.m. ³	
FedEx [®] Envelope up to 8 oz.	FedEx First Overnight [®]	FedEx Priority Overnight [®]	FedEx Standard Overnight [®]	FedEx 2Day [®] A.M.	FedEx 2Day [®]	FedEx Express Saver [®]	
FedEx [®] Pak	\$ 53.15	\$ 23.15	\$ 22.80	\$ 18.13	**	**	
	*	*	*	*	*	*	
Shipments in All Other Packaging / Maximum Weight in Lbs.	1 lb.	\$ 57.69	\$ 27.69	\$ 25.79	\$ 19.13	\$ 17.52	\$ 15.45
	2 lbs.	58.05	28.05	27.24	19.46	17.82	15.73
	3	61.40	31.40	29.61	19.79	18.12	16.18
	4	63.85	33.85	31.93	20.39	18.66	16.46
	5	64.59	34.59	32.50	20.98	19.22	16.74
	6	69.18	39.18	34.30	21.90	20.06	18.09
	7	69.65	39.65	35.40	22.83	20.91	18.38
	8	70.00	40.00	37.26	23.74	21.74	19.11
	9	70.51	40.51	38.58	25.00	22.89	19.26
	10	70.87	40.87	38.73	26.18	23.04	19.45
	11	76.73	46.73	41.52	27.36	25.06	23.10
	12	78.65	48.65	42.91	29.70	26.39	23.60
	13	79.32	49.32	43.38	30.85	28.25	25.73
	14	80.15	50.15	45.06	32.36	29.65	26.01
	15	80.51	50.51	46.80	34.63	30.78	26.30
	16	82.36	52.36	47.78	35.46	31.50	27.76
	17	86.87	56.87	50.21	36.54	32.47	28.66
	18	87.33	57.33	51.71	37.77	33.56	30.46
	19	87.70	57.70	53.10	39.24	34.88	31.24
	20	88.00	58.00	54.08	39.40	36.09	31.53

Figure 3-1: U.S. express package (public) rates example of FedEx.

In addition, Waters Corporation has specific contractual discounts that apply to the basic freight rates (excluding fuel, tax, special handling fees, and other miscellaneous charges). Therefore, the net charge will be less than the published rates.

²For details, please check FedEx shipping rates: <https://www.fedex.com/en-us/shipping/current-rates.html>.

This page was intentionally left blank.

Chapter 4

Baseline Establishment

As the ultimate goal of the team was to optimize package sizes to reduce cost and improve sustainability, it was important to establish a baseline for the empirical packing method currently adopted by GDC. In 2018, there were around 130,000 outbound shipments from GDC, and 56% of them were charged-by-volume domestic FedEx shipments¹. Because a large number of orders was charged by volume, reducing their package sizes yields potential savings. To evaluate the current packing performance, a simple and effective metric was proposed, the average shipping air percentage, \bar{A} , which is defined as follows:

$$\bar{A} = \frac{1}{N_o} \sum_{i=1}^{N_o} a_i \times 100\%$$
$$a_i = 1 - \frac{\sum_{j=1}^{n_i} q_{ij}(l_{ij}w_{ij}h_{ij})}{v_i}$$

where N_o is the total number of charged-by-volume domestic FedEx orders, a_i is the fraction of air in the i^{th} order, v_i is the i^{th} order box volume, n_i is the number of SKU(s) in the i^{th} order, q_{ij} is the quantity of the j^{th} SKU in the i^{th} order, and $l/w/h_{ij}$ are the dimensions of the minimum bounding rectangle (MBR) of the j^{th} SKU in the i^{th} order.

With the average shipping air percentage as a metric, the team was able to not only

¹Including shipments to Canada

understand the current packaging performance but also can now compare the metric before and after a potential solution has been applied to evaluate the effectiveness of the solution.

4.1 Baseline Data

The team acquired all the necessary data to establish the baseline case from Waters SAP. Three primary reports, “SLD Veraction”, “Waters 99% of Demand”, and “GDC_0001”, were downloaded and correlated to create a master order info data sheet shown in the following figure.

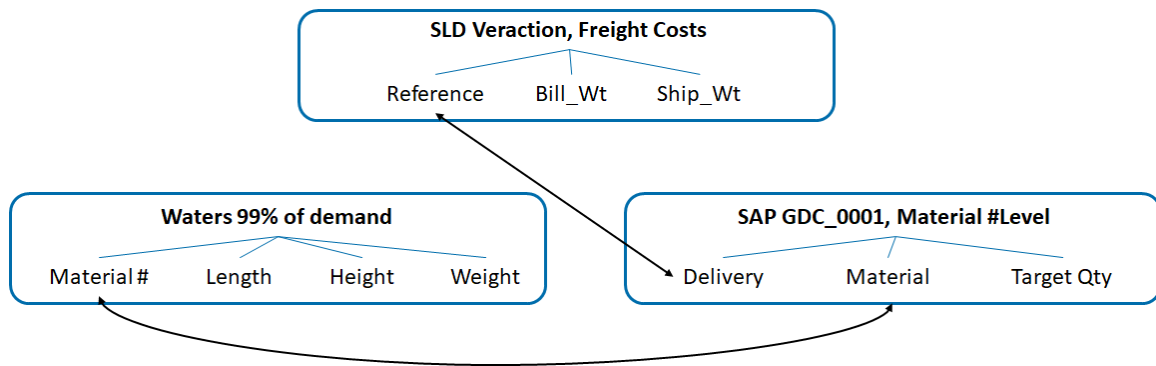


Figure 4-1: Data structure for the GDC packing baseline establishment.

“SLD Veraction” contains high-level order information such as the order reference number, the billable weight, and the ship weight. It is worth mentioning that if an orders ship weight is equal to the billable weight, this order is charged by weight, so reducing the package size of this order will not affect its shipping cost. Therefore, the data used to build the baseline were filtered to only include the orders whose billable weights do not equal to their ship weights, meaning that those orders were charged by volumetric weight.

“GDC_0001” contains detailed SKU information associated with each order, including the delivery² reference number, the material³ reference number and the targeted quantity of an SKU. If an order contains multiple SKUs, this order is listed as

²In Waters, “delivery” is equivalent to “order”.

³In Waters, “material” is equivalent to “SKU”.

multiple lines⁴ in “GDC_0001”, and the number of lines is equal to the number of different SKUs of the order. As shown in the figure above, “GDC_0001” is linked to “SLD Veraction” via the shared order reference numbers.

“Waters 99% of Demand” data sheet is a detailed SKU information look-up table, including data categories such as the material reference number, each orders detailed MBR dimensions and weight measured by an industrial cubic scanner⁵. This database covers over 10,000 SKUs, consisting 99% of the active SKUs. However, due to the early stage of Waters warehouse digitization project, the database is still not comprehensive yet with information for around 4000 SKUs not logged along with some partially missing/fault data associated with the logged SKUs. Therefore, to establish an unbiased baseline, only orders with the complete SKU data were included in the baseline analysis. “Waters 99% of Demand” is linked to “GDC_0001” via the same SKU reference numbers.

4.2 Baseline Assumptions

Due to the incomplete nature of the data, there are several key assumptions for establishing the baseline:

1. Because SAP had not logged the three dimensions of the box used by each order, the team reconstructed the volumes of each box, v_{bi} , from the billable weight in “SLD Veraction” database. Because an orders billable weight is equal to the orders volume divided by a contract dimensional factor, with known billable weights, the volume was estimated to be equal to its billable weight multiplying the dimensional factor. However, such order volume estimation may increase the average shipping air percentage because FedEx rounds each billable volume to the next higher integer. Furthermore, as FedEx dimensions all orders of Waters and transfers order billable weights back to SAP, the team assumed that the dimensioning process of FedEx was fair and accurate.

⁴Rows in database

⁵CUBISCAN 210-DS

2. The team assumed that the data in “Waters 99% Demand” were all accurate and well maintained. However, due to the manual SKU dimensioning process and the existence of SKUs with irregular shapes, a certain percentage of the SKU dimensional and weight data might be off or corrupted.
3. The team also assumed that all orders to be analyzed were overpacked with one or more cartons. However, in reality, there is a small portion of orders that are shipped in own containers (SIOC), such as Lenovo ThinkStation for liquid chromatography equipment. Those SIOC packages cannot be further optimized, but because the SKU data had not contained SIOC tags, the team had no mean to exclude those false opportunities.

4.3 Baseline Analysis Results and Opportunities

There were around 70,000 charged-by-volume FedEx domestic and Canada orders incorporated in the baseline analysis. The average shipping air percentage is 66.7%, with a standard deviation of 20.8%, and lower and upper quantiles of 54.6% and 83.1%, respectively. Based on research conducted by a leading packing solution company, Packsize, for companies that continuously improve their outbound shipping processes, the average shipping air percentage is around 40% to 45% [4]. Therefore, there can be a decent opportunity for Waters to improve its packing operations.

To quantify the potential saving opportunity, several customized box maker services were used to be a benchmark specifying the upper bound of the saving. Those automated box makers can cut and fold the optimal fit corrugated box for each order, which is stacked in a rectangular form, thus minimizing the air volume in packages. A leading custom box maker company claimed that their service could reduce average shipping air percentage down to 18% to 20% [4]. If the orders of Waters are still shipped in corrugated cartons, then other less capital-intensive packaging solutions, such as the software-based ones, can generate a saving within this 40% window, which was regarded as a large saving opportunity that motivated the team and Waters stakeholders to pursue the data-driven packing strategy project further.

Chapter 5

Carton Combination Method

Because a distribution center (DC) has a limited space to hold corrugated carton inventory for packing, there is only a finite number of cartons used. This number usually ranges from 10 to 20 also for keeping the operation lean as packers spend less time on choosing and finding cartons, and the DC wastes less on managing the box inventory. The wasted air volume inside each order is heavily associated with how many distinct sizes of carton a DC has in stock and their dimensions. Generally, the more carton options to choose from, more likely that each order can be packaged with box with minimal wasted air. As for the dimension of each carton options, there always exists an optimal dimensional combination that best fits its unique outbound shipping order dimensional characteristics and thus minimizes the total shipping air volume in all annual orders of a DC, if the number of carton options is predetermined. Therefore, in this chapter, an algorithm is proposed to optimize the dimensions for a set of cartons can be used in a DC, for a specified set size.

5.1 Problem Simplification

To formulate the optimization problem, the following assumptions were made:

1. All orders are only packed with a single carton. If we were to allow multi-carton packing, we would need to specify a rule for deciding whether or not an order needed to be packed into multiple cartons. For example, some potential rule

elements can be “fragile \rightarrow ship in a separate container”, “ship in own container (SIOC) \rightarrow isolate from the rest of SKUs”, “hazardous \rightarrow maximum allowable number of SKUs in a carton is x^1 ”, etc. Thus, to cover those edge cases would require a dramatic increase in the complexity of the implementation. For Waters Corporation, there are only 7% of the total orders shipped with more than one cartons. In addition, to simulate the multi-carton packing, all SKUs should be evaluated and tagged with rule items if necessary, which is hard for most of companies to follow.

2. SKUs of an order can stack on each other without limitations. Similarly, to specify stacking limitations, a complex rule is required to be created and maintained, a status that most companies cannot achieve. Waters actually let packers to decide how to stack SKUs in orders and this practice still fulfills the customers expectations.
3. For a specific company, the annual order characteristics, defined by order’s SKU mix and quantity, can be well predicted by its previous-year order characteristics. Therefore, to optimize the dimension of cartons for a DC, its historical order data serve as inputs.

5.2 Ideal Solver

We assume that we are given annual N lines of order transaction record, $O_1, O_2, O_3, \dots, O_N$. For each order i , we are given its content for the n SKU(s), $S_{i1}, S_{i2}, S_{i3}, \dots, S_{in}$. For each SKU in order i , we are given its quantity, weight and MBR dimensional (length,

¹A positive integer

width, height) data in vector forms:

$$\begin{aligned}
\mathbf{q}(i) &= [q(i, 1), q(i, 2), \dots, q(i, n)] \\
\mathbf{W}(i) &= [W(i, 1), W(i, 2), \dots, W(i, n)] \\
\mathbf{l}(i) &= [l(i, 1), l(i, 2), \dots, l(i, n)] \\
\mathbf{w}(i) &= [w(i, 1), w(i, 2), \dots, w(i, n)] \\
\mathbf{h}(i) &= [h(i, 1), h(i, 2), \dots, h(i, n)]
\end{aligned}$$

The total SKU volume for order i is calculated by²:

$$\begin{aligned}
V(i) &= \mathbf{q}(i) \cdot [\mathbf{l}(i) \odot \mathbf{w}(i) \odot \mathbf{h}(i)] \\
&= q(i, 1)[l(i, 1)w(i, 1)h(i, 1)] + \dots + q(i, n)[l(i, n)w(i, n)h(i, n)]
\end{aligned}$$

And order SKU volume vector is formed by:

$$\mathbf{V} = [V(1), V(2), \dots, V(N)]$$

Similarly, the order weight vector is calculated by:

$$\mathbf{W} = [\mathbf{q}(1) \cdot \mathbf{W}(1), \mathbf{q}(2) \cdot \mathbf{W}(2), \dots, \mathbf{q}(N) \cdot \mathbf{W}(N)]$$

Also, we are given a large pool (quantity of M) of different cartons, C_1, C_2, \dots, C_M , whose dimensional data and weight ratings³ are populated in the following vector

². and \odot are defined as dot and Handamard product, respectively.

³The maximum allowable weight to be packed

forms⁴:

$$\begin{aligned}
\mathbf{l}_c &= [l_{c_1}, l_{c_2}, l_{c_3}, \dots, l_{c_M}] \\
\mathbf{w}_c &= [w_{c_1}, w_{c_2}, w_{c_3}, \dots, w_{c_M}] \\
\mathbf{h}_c &= [h_{c_1}, h_{c_2}, h_{c_3}, \dots, h_{c_M}] \\
\mathbf{Wr}_c &= [Wr_{c_1}, Wr_{c_2}, Wr_{c_3}, \dots, Wr_{c_M}]
\end{aligned}$$

The box volume vector is calculated by:

$$\begin{aligned}
\mathbf{V}_c &= \mathbf{l}_c \odot \mathbf{w}_c \odot \mathbf{h}_c \\
&= [l_{c_1}w_{c_1}h_{c_1}, l_{c_2}w_{c_2}h_{c_2}, \dots, l_{c_M}w_{c_M}h_{c_M}]
\end{aligned}$$

In addition, we are given a a pre-determined number of distinct carton options to be used in a DC, m , and $m \ll M$.

In order to determine the optimal carton combination based on the above-given inputs, a mixed integer linear programming (MILP) problem was formed as follows:

- *Variables:*
 1. Define an $N_o \times N_b$ box decision (binary) matrix, \mathbf{B}_d , and each element is either 0 or 1. This variable denotes which box is assigned to each order.
 2. Define an $1 \times N_b$ box activation (binary) vector, \mathbf{B}_a , and each element is either 0 or 1. This variable denotes whether or not a box is chosen or activated.
- *Objective:* Similar to the baseline case, the cost function is defined as the total air volume inside all packed cartons⁵:

$$V_a = \|\mathbf{B}_d \mathbf{V}_c^T - \mathbf{V}^T\|_1$$

The objective of the optimization is to minimize the total shipping air volume.

⁴The subscript c denotes that a variable is associated with cartons.

⁵Use Euclidean 1-norm to represent sum of vector elements.

- *Constraints:*

1. As stated in 5.1, we assume that each order will be packed into a single carton; thus the sum of each row of the box decision matrix is equal to 1:

$$\mathbf{B}_d [1]_{M \times 1} = [1]_{N \times 1}$$

2. The volume of each order should be less than the volume of the chosen box for that order:

$$\mathbf{B}_d \mathbf{V}_c^T - \mathbf{V}^T \geq [0]_{N \times 1}$$

3. The weight of each order should be less than the weight rating of the chosen box for that order:

$$\mathbf{B}_d \mathbf{W} \mathbf{r}_c^T - \mathbf{W}^T \geq [0]_{N \times 1}$$

4. In order to assure that at most m boxes are activated and that each order is assigned to an active box, the matrix \mathbf{B}_d should only have m columns containing 1's. This was realized by the following two constraints:

$$\begin{aligned} \begin{bmatrix} - & \mathbf{B}_a & - \\ & \vdots & \\ - & \mathbf{B}_a & - \end{bmatrix}_N - \mathbf{B}_d \geq [0]_{N \times M} \\ \|\mathbf{B}_a\|_1 = m \end{aligned}$$

5. Lastly, there should exist a SKU 3-D stack for each order such at the dimension of the stacks MBR is less than that of the chosen box. This constraint will be discussed in greater details in 5.2.1.

5.2.1 SKU Dimensional Constraint

Only relying on the order volume constraint (5.2—item 2 of *Constraints*) is not enough to reflect the real packing situation. For example, if an incoming order has a very long SKU with only a tiny cross-sectional area, whose volume is still small, then based on the order volume constraint, the optimization program will tend to choose a box whose volume is just a bit larger than the SKU volume. However, for such SKU with a large dimensional aspect ratio, its longest dimension may exceed the largest dimension of the chosen box, which means that this SKU cannot be fit into the chosen box. Therefore, in order to optimize the carton combination practically, a constraint, which compares the box dimensional data with the SKU dimensional data, is required.

For generality, it would be ideal to propose an algorithm that can stack the SKUs in each order into a shape that is “closest” to a rectangular and then output the three dimensions of the MBR of that SKU stack. Here, the most difficult challenge is to mathematically define what “a SKU stack closest to a rectangular” means. Based on literature surveys, solutions to the 3-D bin packing problem might provide a promising angle of attack.

According to Wolfram MathWorld, the general bin-packing problem is defined as: “it is the problem of packing a set of items into a number of bins such that the total weight, volume, etc. does not exceed some maximum value [1].” Its commonly referred definition is more narrow: “In the bin packing problem, items of different volumes must be packed into a finite number of bins or containers each of volume, V , in a way that minimizes the number of bins used [2].” The key information extracted from those definitions is that solutions to the bin-packing problem always include strategies to fit items (SKUs) into boxes. There have been multiple packing algorithms in existence such as the First Fit method, the First Fit Decreasing method, and the Last Fit method [7]. However, those packing methods do not always guarantee an optimal packing result and are subjected to lower bounds of the void volume in a package. In addition, most of the literature is limited to one-dimensional objects. However, one of

the most advanced packing algorithms has been presented by Dube and Kanavathy on their 2006 paper of Optimizing Three-dimensional Bin Packing Through Simulation [7] that solves the above issues. The superiority of their algorithm is shown in the following three aspects:

1. The algorithm is specifically designed for the **3-D** bin packing problem.
2. The algorithm uses the Best Fit method, allowing orthogonal rotation of items, which leads to the true optimal packing, and always guarantees a solution.
3. The solution of the algorithm is computationally feasible to obtain.

Therefore, for our application, it might be possible to modify their 3-D bin packing algorithm into 3-D packing algorithm, which is the simplest special case of the 3-D bin packing problem, solving for the optimal packing strategy for only a single box. However, as further research was conducted, because of the the 3-D packing algorithm solves heuristically, incorporating Dube and Kanavathy’s algorithm into the SKU dimensional constraint transfers the optimization into Mixed Integer Dynamic Programming⁶, which would steeply increase the implementation challenge and also potentially make the problem computationally infeasible. Therefore, a simpler alternative SKU dimensional constraint should be proposed to efficiently approximate the constraint built upon the 3-D packing algorithm.

On average, in 2018, each order of Waters contains only 2.2 SKUs. Therefore, a naive assumption regarding the comparison between SKU and carton dimensions was proposed that if the largest SKU of an order can be fit in its carton, the entire order is assumed to be able to be fit in that carton. More specifically, the largest SKU dimension of an order is constrained to be smaller than the largest chosen-carton dimension, the second-largest SKU dimension is constrained to be smaller than the second-largest carton dimension, and the third largest SKU dimension is constrained to be smaller than the third largest carton dimension. This naive assumption is likely to work best for orders with a small number of SKUs. So, since the average

⁶Lose both linear and static features from the original setup.

number of SKU per order was quite small for Waters Corporation, such simplified SKU dimensional constraint should be able to simulate the real packing situation and help deliver a meaningful optimized carton combination result. However, to fully validate the assumption, we need to sample at least a week's worth of orders to see if most of them can fit or not.

Before specifying the simplified SKU dimensional constraint, the input data discussed in 5.2 were further calculated as follows.

For any p by q matrix \mathbf{S} , define $m(\mathbf{S}, a)$ as the a^{th} largest element in each column of \mathbf{S} , and so it returns a row vector with a dimension of q .

We expand the box dimensional data into a matrix, \mathbf{D}_c :

$$\mathbf{D}_c = \begin{bmatrix} - & \mathbf{l}_c & - \\ - & \mathbf{w}_c & - \\ - & \mathbf{h}_c & - \end{bmatrix}$$

The largest, second-largest and third largest box dimension vectors are:

$$\mathbf{Bd}_1 = m(\mathbf{D}_c, 1)$$

$$\mathbf{Bd}_2 = m(\mathbf{D}_c, 2)$$

$$\mathbf{Bd}_3 = m(\mathbf{D}_c, 3)$$

The largest SKU length, width and height vectors for all orders are:

$$\mathbf{L}_s = [\max(\mathbf{l}(1)), \max(\mathbf{l}(2)), \dots, \max(\mathbf{l}(N))]$$

$$\mathbf{W}_s = [\max(\mathbf{w}(1)), \max(\mathbf{w}(2)), \dots, \max(\mathbf{w}(N))]$$

$$\mathbf{H}_s = [\max(\mathbf{h}(1)), \max(\mathbf{h}(2)), \dots, \max(\mathbf{h}(N))]$$

We expand the largest SKU dimensional data into a matrix, \mathbf{D}_s :

$$\mathbf{D}_s = \begin{bmatrix} - & \mathbf{L}_s & - \\ - & \mathbf{W}_s & - \\ - & \mathbf{H}_s & - \end{bmatrix}$$

The largest, second-largest, and third largest SKU dimension vectors for all orders are:

$$\mathbf{Sd}_1 = m(\mathbf{D}_s, 1)$$

$$\mathbf{Sd}_2 = m(\mathbf{D}_s, 2)$$

$$\mathbf{Sd}_3 = m(\mathbf{D}_s, 3)$$

Finally, the simplified SKU dimensional constraint is expressed as the following three inequalities:

$$\mathbf{B}_d \mathbf{B} \mathbf{d}_1^T - \mathbf{Sd}_1^T \geq [0]_{N \times 1}$$

$$\mathbf{B}_d \mathbf{B} \mathbf{d}_2^T - \mathbf{Sd}_2^T \geq [0]_{N \times 1}$$

$$\mathbf{B}_d \mathbf{B} \mathbf{d}_3^T - \mathbf{Sd}_3^T \geq [0]_{N \times 1}$$

Up to this point, the optimization problem was completely set up.

5.2.2 Data Input

The carton pool, C_1, C_2, \dots, C_M , was established based on all available corrugated boxes on “Packaging Supplies.com”. It was assumed to be the only carton supplier and all the corrugated boxes would be readily available with an unlimited inventory. There are total of 1032 different cartons available, so M is equal to 1032. It was also assumed that the 1032 options were comprehensive enough to cover most of standard cartons available on the market. Each carton is with its weight rating and three dimension information, which was input in the corresponding vectors, $\mathbf{W}r_c, l_c, w_c$,

and h_c . The box information was manually collected from the website interface whose screen-shot is shown as follows [5].

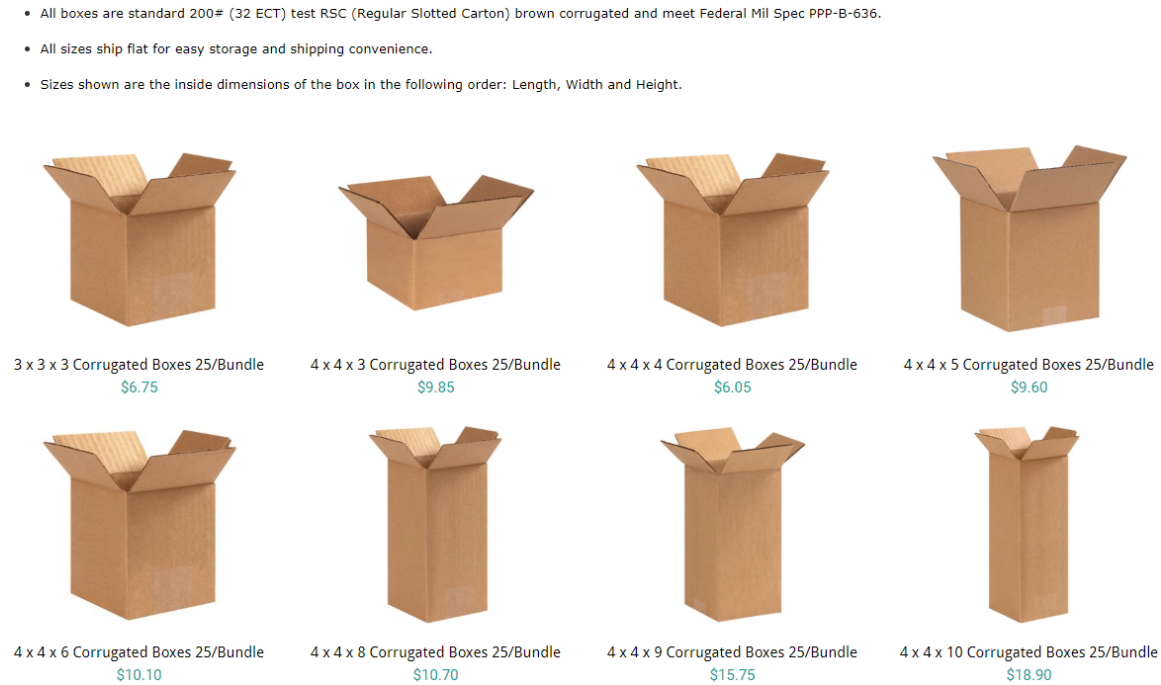


Figure 5-1: Screen-shot of box info from Packaging Supplies.com.

The carton combination method shares the same data used to establish the baseline case as discussed in 4.1. By linking the three databases (“SLD Veraction”, “Waters 99% of Demands”, and “GDC_0001”), a master order databases was created whose example is shown in the following table⁷.

Delivery	Material	Target Quantity	L (<i>in</i>)	W (<i>in</i>)	H (<i>in</i>)	Weight (<i>lbs</i>)
0646	WAT951	20	6.6	5.3	2.8	3
0647	7781	1	4	5.5	1.9	1.2
0656	WAT388	3	3.7	3	0.7	2.3
0656	WAT344	1	3.1	9.3	2	5

Table 5.1: Example of master order info database.

It is worth noting that there are 2 lines for the order 0656, which means that the order 0656 contains two SKUs, namely WAT388 and WAT344. So, for this 3rd order,

⁷Table data were modified to protect Waters propriety information.

n_s is equal to 2, and the order data vectors are shown as follows:

$$\begin{aligned}\mathbf{q}(3) &= [3, 1] \\ \mathbf{W}(3) &= [2.3, 5] \\ \mathbf{l}(3) &= [3.7, 3.1] \\ \mathbf{w}(3) &= [3, 9.3] \\ \mathbf{h}(3) &= [0.7, 2]\end{aligned}$$

All other order data was input following this format.

However, for practicality, orders that match with any of the following three conditions were not included specifically for the carton combination method:

1. In Evaluation of Outbound Operations Improvement Projects for Distribution Centers [10], Yu proposed a packing strategy for small orders with SKU volumes less than 280 in^3 . His method can ensure all small orders to be charged by weight, a situation that is equivalent to have zero air volume in all small orders. Therefore, for orders with volumes less than 280 in^3 , his method reaches the theoretical optimum, thus no optimization further being required.
2. For Waters' GDC, there are two types of SKUs that needs special packing handling: the SKUs that contain hazardous materials or that require refrigeration. Orders including those two types of SKUs were excluded from the master order database as they can only be packed within a single type of certified carton.
3. The largest 5% orders that only contain a single SKU were assumed to be ship in own container (SIOC). The package of those SIOC orders cannot be further optimized as the original packages have to be intact. Therefore, all SIOC orders were excluded from the optimization. This assumption was validated by the GDC manager of Waters Corporation.

After excluding the above three types or special orders, around 47,000 orders were input into the optimization program. Based on the baseline calculation discussed in

chapter four, a new baseline was established for those 47,000 orders. The updated annual average shipping air percentage of them is 60.3%, with a standard deviation of 21.2%, lower and upper quantiles of 47.6% and 75.6%, respectively. This updated baseline was utilized to evaluate the performance of the carton combination method. In order to have a fair comparison with the updated baseline, the number of carton options, m , was set to be the same as that of GDC, which is 12.

5.2.3 Ideal Solver Results and Challenges

The optimization problem presented in 5.2 was implemented via MATLAB R2019a and the script is shown in Appendix A, Listing A.1. The standard optimization toolbox of MATLAB R2019a was used to facilitate the implementation of an efficient MILP solver. The optimized combination of 12 cartons is summarized in the following table:

New Box Types	Dimensions (<i>in</i>)		
	L	W	H
Type1	36	24	12
Type2	26	22	12
Type3	36	24	10
Type4	26	16	16
Type5	22	16	16
Type6	26	15	12
Type7	26	16	10
Type8	36	24	4
Type9	22	22	6
Type10	14	8	5
Type11	14	10	2
Type12	12	6	2

Table 5.2: Optimized carton options via ideal solver.

After optimization, the new average shipping air percentage is 51.2%, resulting in a reduction of 9.1% from the baseline of 60.3% air. However, the optimization result presented was only based on inputting the first 2000 orders, instead of the total of 47,000 orders; the optimized result from such a small number of order inputs cannot conclude the performance of the carton combination method as the first 2000 orders,

consisting only 4.3% of the total orders, are not able to fully represent the dimensional characteristics of the annual orders. It took more than 5 hours to optimize the first 2000 orders⁸ and for the fact that MILP is NP-hard, optimizing all 47,000 orders using the current problem formulation might make the solver computationally infeasible. Therefore, an easier alternative optimization formulation was required to reasonably approximate the ideal formulation, along with further input data simplification.

5.3 Quick Solver

5.3.1 Box Pool Down-selection

To improve the efficiency of the algorithm, one potential route is to reduce the size of the box pool so as to narrow down the search space for the carton combination method. Since a preliminary set of cartons was achieved by running the algorithm with the first 2,000 orders, we assumed that the carton dimensions, concluded by the global optimization, would be close to those of the preliminary set. Therefore, the box pool, C_1, C_2, \dots, C_M , was first sorted, then for each carton in the preliminary set, the additional nine closest larger and smaller⁹ cartons were also chosen, along with the cartons in the preliminary set, to form a downsized carton pool for the quicker algorithm. Also, the boxes with the same three dimensional values but in different length, width and height permutations were treated to be the same, and all duplicates were removed. After the box down-selection process, the pool size was reduced from 1032 SKUs to 210.

5.3.2 Constraint Relaxation

We found that the constraint specifying the allowable number of carton options in a DC (5.2—item 4 of *Constraints*) was the main source causing the slow optimization program. It was reasonable to further investigate the optimized result by relaxing this constraint. In other words, the relaxed optimization allows an order to be freely

⁸Via an Alianware M17

⁹Additional 18 cartons were chosen for each carton in the preliminary set.

matched with the best-fit carton within the down-sized pool, ensuring EACH order’s package to contain the smallest wasted air volume. Furthermore, after the constraint relaxation, the box matching problem can be more efficiently solved by an iterative solver discussed in 6.1, instead of the MILP solver. By inputting the down-selected box pool to the iterative solver, the result is shown in the following figure.

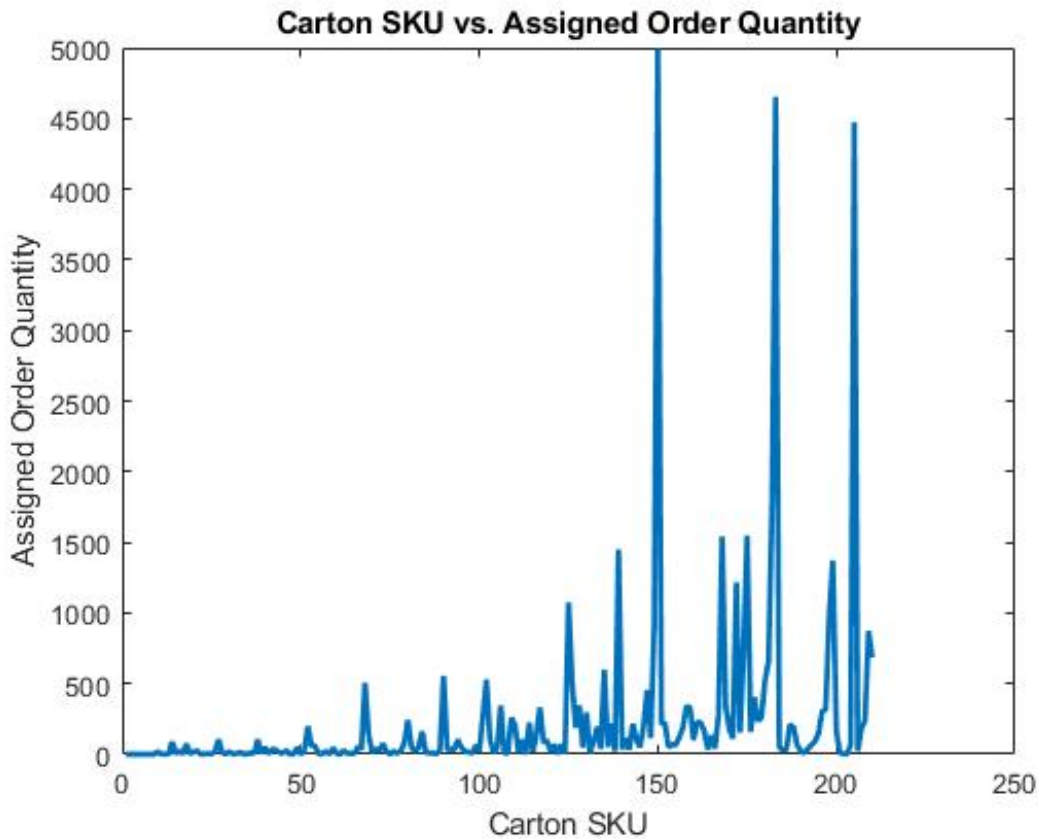


Figure 5-2: Order quantities assigned to each carton in pool.

It is worth noting that, for the x-axis of the above figure, the carton options were sorted in descending order with the box volume as the first priority, then the largest dimension, then the second-largest dimension, and finally the third largest dimension. The results turned out to be quite interesting, with several significant peaks that are associated to the most popular cartons. It was tempting to directly select the most popular 12 carton options to form a potential solution, but those 12 cartons cannot cover orders with a volume larger than 2880 in^3 , and thus results in an impractical carton combination. Therefore, a systematical mean to down-select m cartons to

form the combination based on those peaks is required to sub-optimally minimize the annual average shipping air percentage.

5.3.3 Carton Worth Down-selection Method

The carton set selection criteria should be biased towards larger boxes, because if more efforts are spent on ensuring larger orders to have a better fit, more significant air volume reduction will be generated. In other words, choosing cartons that better fit larger orders creates a higher return of investment (ROI). Since the carton pool had been sorted in descending order by volume, the carton worth for each carton is defined as:

$$Wor_c = \frac{N_{oa}}{I_c}$$

where Wor_c is the worth of a carton, N_{oa} is the assigned order quantity to that carton, and I_c is the carton index assigned to each carton based on the descending sorting rule discussed in 5.3.1. After the worth of each carton was calculated, the cartons with the highest 12 worth's were chosen to form the proposed combination. We used this carton combination to re-run the optimization, and we found that the new average shipping air percentage was reduced to 45%, and compared with the updated baseline, a 15% air volume reduction was realized.

To evaluate why the carton worth down-selection method works, k -means analysis was conducted on a 1-D array, whose elements contain the assigned order quantities to the 210 cartons. k -means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells [3]. For the 1-D k -mean clustering, the result is a partitioning of the data into k segments on an ordered number line.

By applying k -means clustering to the 1-D array mentioned above, the assigned order quantities to those cartons were clustered into 12 groups. In other words, the figure above were sliced into 12 layers horizontally based on the height (y-coordinate)

of data points. The k -means results computed via MATLAB R2019a is shown in the following figure.

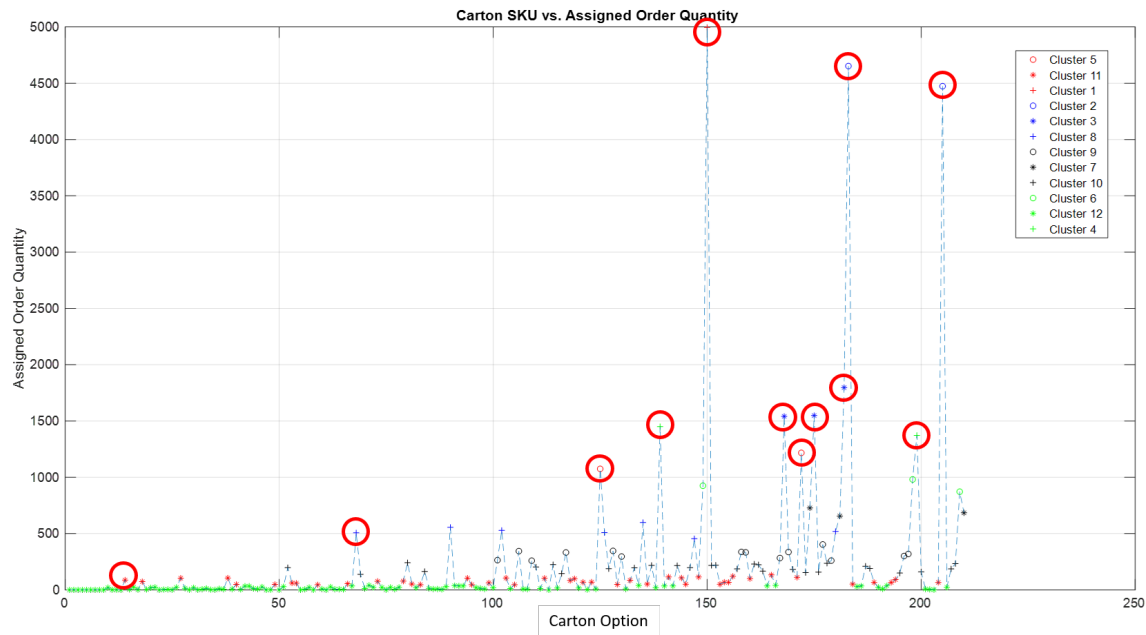


Figure 5-3: Order quantity assigned to each carton in the pool with the k -means clustering and carton combination results called out.

The k -mean clustering result further supports the effectiveness of the carton worth down-selection method. The first 5 clusters associated with the most popular cartons were fully covered as they have higher ROI's. Then, because of the diminishing returns for further guaranteeing a good fit for the remaining small orders, 6th and 7th clusters, which are related to small volume orders, were totally skipped. As expected, the LARGEST cartons that belong to the two clusters (8th and 11th) associated with the larger orders were also chosen to ensure all orders can be covered by the improved carton SKU combination. In addition, the fact that even though the those two types of large cartons are only perfect for a small amount of orders, they were still chosen, indicates the method is properly biased to select cartons that can better accommodate larger orders.

The proposed set of 12 cartons is summarized in the following table:

Box Types	Dimensions (<i>in</i>)		
	L	W	H
Type1	36	24	20
Type2	26	16	16
Type3	20	12	12
Type4	17	12	10
Type5	18	12	7
Type6	20	12	3
Type7	12	10	6
Type8	24	6	4
Type9	15	12	3
Type10	11	7	7
Type11	12	10	2
Type12	12	6	2

Table 5.3: Carton SKU combination via quick solver—set of 12 cartons.

5.3.4 Results and Discussion

The carton combination method was also run for the case where m is equal to 20. This time, the average shipping air percentage was further reduced down to 40%. However, the additional 5% air volume reduction is achieved with the cost of increasing the difficulty of packing operations, as there will be more carton options to manage, and packers will spend more time to decide and find a carton for each order. The proposed set of 20 cartons is summarized in the following table:

Box Types	Dimensions (<i>in</i>)		
	L	W	H
Type1	36	24	20
Type2	40	20	20
Type3	26	16	16
Type4	26	15	12
Type5	26	20	8
Type6	20	12	12
Type7	26	20	4
Type8	17	12	10
Type9	13	13	9
Type10	18	12	7
Type11	20	12	3
Type12	12	10	6
Type13	24	6	4
Type14	18	16	2
Type15	15	12	3
Type16	11	7	7
Type17	12	10	2
Type18	12	5	4
Type19	12	6	2
Type20	8	8	2

Table 5.4: Carton options via quick solver—set of 20 cartons.

Generally, the more carton types that a set contains, the less average shipping air percentage is, but the trend is subject to the diminishing return effect, and the air percentage is minimized when having all the cartons active, as shown in following figure.

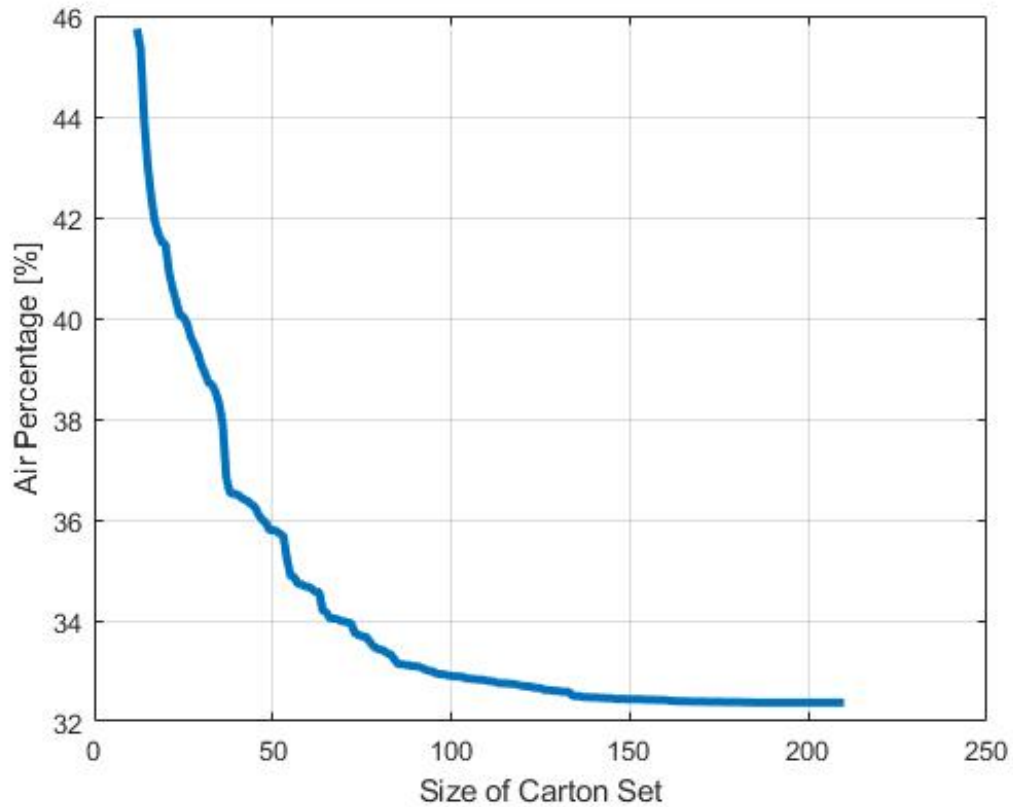


Figure 5-4: Averaged shipping air percentage vs. size of carton set.

Therefore, for a DC, a financially-oriented metric should be defined in order to optimize the trade-off between the freight saving generated by having a large set of cartons versus the increased operational costs. Such a metric is tailored to each company's unique status and needs.

This page was intentionally left blank.

Chapter 6

Carton Selection Algorithm

In chapter 5, a method was discussed to reduce the air volume in shipments by improving the dimensions for a set of cartons used in a DC. To achieve the targeted wasted air reduction, it is important to develop an algorithm to inform packers, in real-time, which carton in the set can be the best fit with an incoming order.

6.1 Algorithm Overview

For Waters' GDC, as discussed in 2.3.5 incoming orders are pushed periodically from ERP to GDC's WMS. Therefore, an algorithm, which constantly runs in the background to match cartons from the improved carton SKU set to incoming orders, is desired. This algorithm is named as Carton Selection Algorithm whose flow diagram is shown in the following figure:

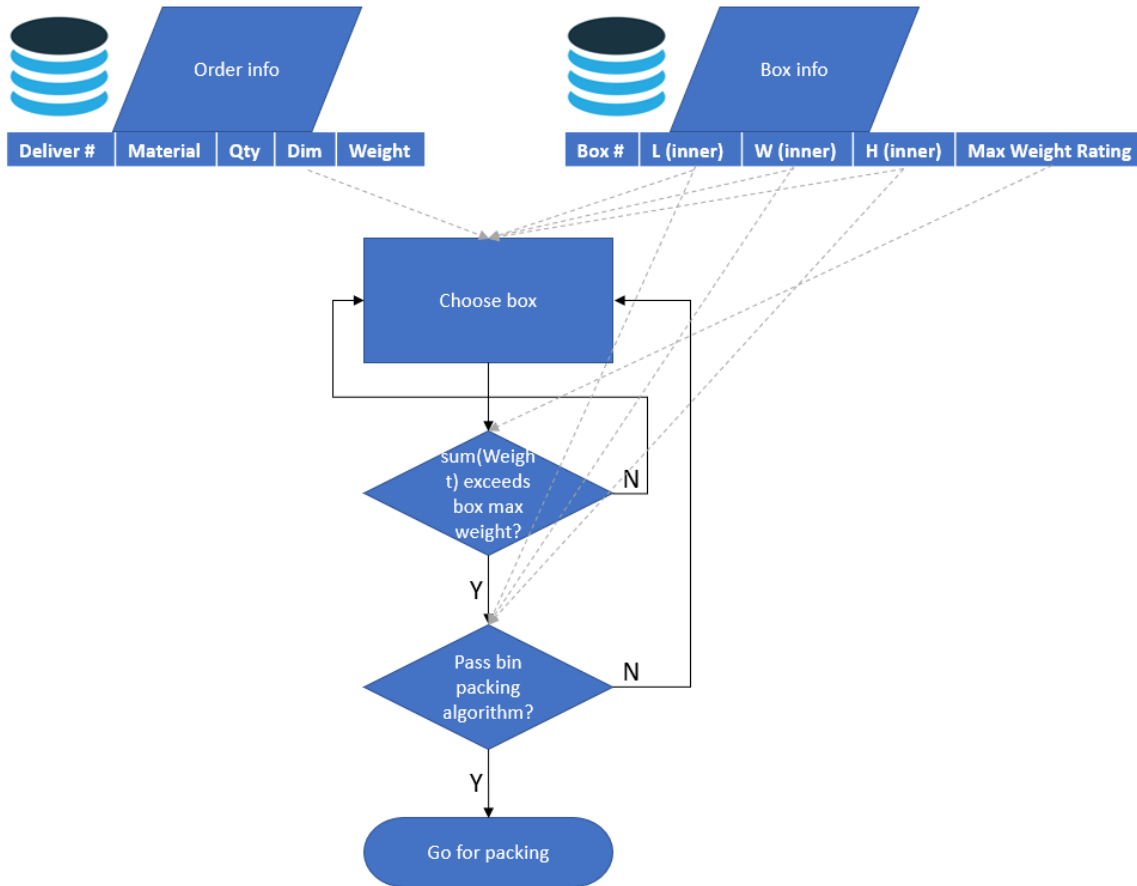


Figure 6-1: Carton selection algorithm flow diagram.

- *Inputs:* The order data input is exactly the same as the inputs used by the Carton Combination method as mentioned in 5.2.2. The improved combination of m cartons SKUs will be inputted as the box info data. It is worth mentioning that the box data will be sorted in an ascending order with the box volume as the first sorting priority, then the largest dimension, then the second-largest dimension, and finally the smallest dimension to break up any tie in previous level of sorting.
- *Box Choosing:* Because the number of distinct box SKUs in an improved set is usually only ranging from 10 to 20 for a DC, which is a small number, the box choosing process can run iterative from the smallest box (the first in the box data) to the largest (the last in the box data). During an iteration, a box is picked to verify whether the incoming order can be fit in the carton, and the

order's weight is not exceeding the box's weight rating. If any of the above two criteria cannot be satisfied, the next box is chosen to compare with the order in the same manner, and this process keeps going until the smallest suitable box is found for the order.

- *Weight Verification:* During an iteration of the box choosing process, the total weight of an order is compared with the weight rating of the chosen box. If the total weight of an order exceeds the weight rating of the box, another box choosing iteration starts.
- *SKU Dimension Verification:* The SKU dimension verification uses the same concept for establishing the SKU dimensional constraint as discussed in 5.2.1. The largest SKU dimension of an incoming order is compared with the largest dimension of the chosen box, the second-largest SKU dimension is compared with the second-largest box dimension, and the third-largest SKU dimension is compared with the third-largest box dimension. During an iteration, if all three SKU dimensions are all smaller than those of the box picked for this iteration, the picked box is the smallest box that ensures a fit of the order.

Such a light algorithm has the potential to be easily integrate with the WMS system of Waters Corporation. The algorithm was prototyped via MATLAB R2019a whose script is attached at Appendix A, Listing A.2.

6.2 Updated Packing SOP

Because of the introduction of the Carton Selection algorithm into the Waters' WMS, the SOP for GDC packing station was updated as follows:

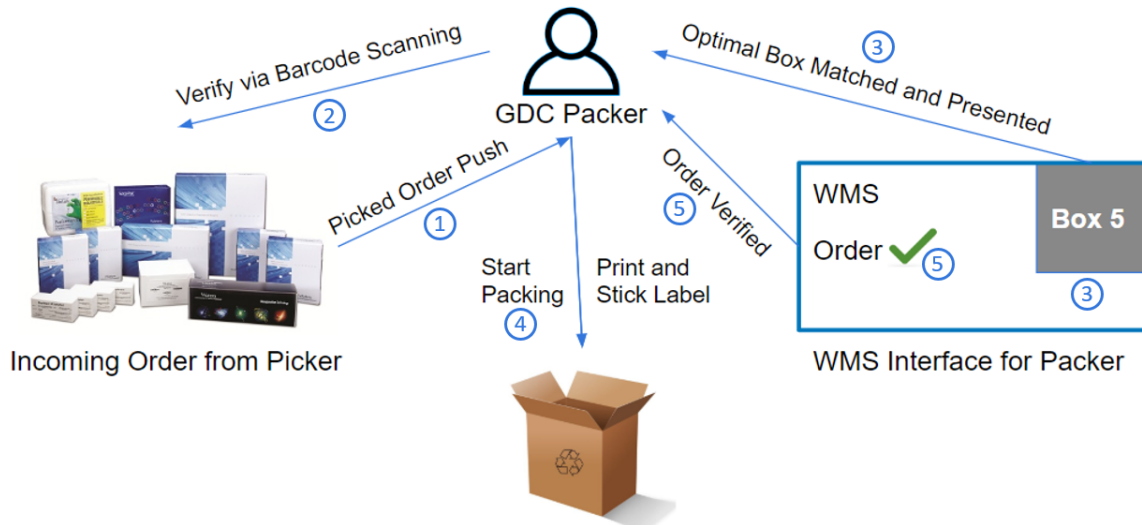


Figure 6-2: New GDC packing SOP.

1. An order picker pushes a cart with full of SKUs for multiple orders. A packer selects one order to pack.
2. The packer scans the barcode for the order to retrieve all order data info from SAP.
3. In the meantime of step 2, once the order data is loaded, the Carton Selection algorithm is running in the background to find the best fit carton for this order and clearly displays the box type on the interface of WMS.
4. After the packer sees the recommended box info, the packer picks the corresponding box and proceeds to scan and pack SKUs. If the recommended box cannot successfully fit all SKUs for this order, the packer selects another box based on experience and proceeds; the packer is prepared to manually input the final box used into WMS, after the order is verified. Such outlier data are constantly collected and will be analyzed periodically to continuously improve the algorithm.
5. Once all SKUs are scanned and packed, an “order verified” signal is clearly visualized on WMS, and the packer is ready to print the shipping label and seal the box.

6. The packer proceeds to the next order packing process.

To validate the above proposed SOP, a three-stage adoption plan was recommended to Waters for future implementation:

1. *Pilot Stage:* The SOP is only adopted to orders with one or two SKU(s), and all the other orders are packed with the original packing processes. The reason why using those orders in the pilot project is to preliminarily validate the naive assumption mentioned in 5.2.1 to estimate if an order can be packed by a chosen carton, based on the fact that Waters' orders contain 2.2 SKUs on average. If there is a significant amount of one-SKU or two-SKU orders cannot be fit in the cartons selected by the algorithm, then that naive assumption, the foundation of the Carton Selection algorithm, needs to be reconsidered. Alternatively, if most of those orders can be packed without a problem, the adoption plan can move to the next stage.
2. *Ramp-up Stage:* At this stage, orders with more SKUs are gradually allowed to be packed with the carton recommendations from the Carton Selection algorithm. Companies are recommended to expand the adoption step by step, and during each step, orders with y^1 more SKUs are allowed to be packed with the new SOP. The step size depends on how aggressively companies want to improve their packing operations. During each step, orders that fail to fit in their recommended cartons are recorded for further investigation whose results are used to improve the algorithm; once the algorithm can correctly cover over 95% orders of a week, companies can take the next adoption step. After all the steps during the ramp-up stage are accomplished, the carton selection algorithm is expected to recommend the best-fit cartons for over 95% orders with less than or equal to 10 SKUs.
3. *Full Adoption Stage:* At this stage, most of the orders are packed with the new SOP, and packers should be familiar with the Carton Selection algorithm. The

¹A positive integer

remaining task is to conduct continuous improvement for the new SOP, and also maintain the SKU database regularly to ensure it is accurate and up-to-date.

Chapter 7

Conclusion and Future Works

This thesis proposed and presented prototypes of the Carton Combination method and the Carton Selection algorithm to tackle the overpacking problem in the global distribution center of Waters Corporation. The Carton Combination leverages the mixed integer linear programming with the constraint relaxation feature and the Metric-based Carton Down-selection method to find a near-optimal solution for the combination of carton SKUs to be used in a DC from a large pool of carton SKU selections. The Carton Selection algorithm uses a light iterative solver to match each incoming order with the best fit box from the optimized box SKU combination. By combining both solutions an improved packing strategy is formed and is ready to be tailored to many other distribution centers, warehouse and fulfillment centers. My colleague, Yu, had conducted a financial evaluation of this packing strategy, and an annual freight saving over 83,000 USD was projected because of the reduction of the void inside order packages. Moreover, around 170,000 ft^2 of the corrugate cardboard reduction were also predicted as the bonus sustainability saving, resulting about 20 ton CO₂ emission reduction per year [10].

However, the Carton Combination method can be further improved from the following aspects:

- As mentioned in 5.2.2, if the order info data can be tagged with the appropriate rules, it is possible for the method to incorporate the multi-carton packing

scenario.

- More historical data may even increase the performance of the method, so it has potentials to input more than one year order data to optimize, if more historical data are readily accessible.
- It might be possible, by tweaking the ideal MILP problem formulation, to create an improved optimization problem with a strong constraint relaxation capability, so that the Carton Combination method can be directly solved by a MILP solver within a short period of time rather than implementing those approximation alternatives which can only reach a sub-optimal solution.

To put the Carton Selection algorithm onto the production environment of SAP, the following suggestions were made:

- Since the SKU dimensional and weight data are the foundation of the Packing Strategy, their fidelity greatly affects the performance. Therefore, it is important for a DC to develop a long-term strategy to contiguously maintain and improve the SKU data.
- In addition, it would be nicer if a SKU data pipeline can be built among GDC, its suppliers, vendors and manufactures, so that any changes of SKU packages can be immediately reflected on the master SKU database.
- As mentioned in 6.2 if both tags used for SKU separation and stack limitation can be created, the algorithm will encounter less edge cases where manual packing is needed. Therefore, it is desired to have a long-term data logging planning for creating and maintaining SKU tags for special packing requirements.
- Another pilot project about optimizing the package of each individual SKU from the manufacturer is also desired, as the project can further reduce shipping-related cost. It is worth noting that the mechanism to optimize the SKU own packages is quite similar to that of the Carton Combination method.

- It is desired to have a strong commitment from the IT department of a DC to seamlessly integrate the Packing Strategy and its continuous improvement program to the real production environment of SAP.

This page was intentionally left blank.

Appendix A

MATLAB Script

Listing A.1: Carton Combination Method Prototype

```
clear all
close all
clc

%load order data
range = 47129;

opts = detectImportOptions('20190802_Order_Vol.xlsx');
opts.DataRange = ['1:',num2str(range)];
preview('20190802_Order_Vol.xlsx',opts)
order_dim = readmatrix('20190802_Order_Vol.xlsx',opts);

dims1 = detectImportOptions('20190802_SKU_Dim_Max_1.xlsx')
;
dims1.DataRange = ['1:',num2str(range)];
preview('20190802_SKU_Dim_Max_1.xlsx',dims1)
order_dim_max1 = readmatrix('20190802_SKU_Dim_Max_1.xlsx',
    dims1);

dims2 = detectImportOptions('20190802_SKU_Dim_Max_1.xlsx')
;
dims2.DataRange = ['1:',num2str(range)];
preview('20190802_SKU_Dim_Max_2.xlsx',dims2)
order_dim_max2 = readmatrix('20190802_SKU_Dim_Max_2.xlsx',
    dims2);

dims3 = detectImportOptions('20190802_SKU_Dim_Max_3.xlsx')
;
```

```

dims3.DataRange = ['1:',num2str(range)];
preview('20190802_SKU_Dim_Max_3.xlsx',dims3)
order_dim_max3 = readmatrix('20190802_SKU_Dim_Max_3.xlsx',
    dims3);

refs = detectImportOptions('20190802_Order_Reference.xlsx'
    );
refs.DataRange = ['1:',num2str(range)];
preview('20190802_Order_Reference.xlsx',refs)
order_ref = readmatrix('20190802_Order_Reference.xlsx',
    refs);
order_ref = str2num(cell2mat(order_ref));

%load box data
cartons = detectImportOptions('20190731_Box_Data_200.xlsx'
    );
cartons.DataRange = '1:210';
preview('20190731_Box_Data_200.xlsx',cartons)
box_dim = readmatrix('20190731_Box_Data_200.xlsx',cartons)
    ;

% Concatenate order data
%preprocessing order data: eliminate FED Box and E-C
order_info = [order_ref, order_dim, order_dim_max1,
    order_dim_max2, order_dim_max3];
idx = any ((order_info(:,2) >= 20800) | (order_info(:,3)
    >=40) | (order_info(:,4)>=26)|(order_info(:,5)>=20) ,2)
    ;
out = order_info(idx,:);
order_info(idx,:) = [];

order_dim = order_info(:,2);
order_dim_max1 = order_info(:,3);
order_dim_max2 = order_info(:,4);
order_dim_max3 = order_info(:,5);

box_name_neu = 1:length(box_dim);
box_name = string(box_name_neu);

box_vol = box_dim(:,4);
box_dim_max1 = box_dim(:,1);
box_dim_max2 = box_dim(:,2);
box_dim_max3 = box_dim(:,3);

test_order_num = length(order_dim);

```

```

% num_box_mix = 14;

boxprob = optimproblem;

%define LP var.
box = optimvar('box',test_order_num,box_name,'Type','integer',
    'LowerBound',0,'UpperBound',1)

% box1 = optimvar('box1',1,box_name,'Type','integer',
    'LowerBound',0,'UpperBound',1)

options = optimoptions('intlinprog','IntegerTolerance',1e
    -5,'Display','iter','PlotFcn',@optimplotmilp,...
    'MaxTime',43200);

%define LP obj.

air_vol = sum(box*box_vol-order_dim);
boxprob.Objective = air_vol;

%define cons.
single_box_pk = box*ones([length(box_name),1]);

% lim_box_mix = sum(box1);
% lim_box_free = repmat(box1,range,1) - box;

box_larger = box*box_vol-order_dim;
dim_larger1 = box*box_dim_max1 - order_dim_max1;
dim_larger2 = box*box_dim_max2 - order_dim_max2;
dim_larger3 = box*box_dim_max3 - order_dim_max3;

boxprob.Constraints.consSBP = single_box_pk == 1;

% boxprob.Constraints.consLBM = lim_box_mix == num_box_mix
;
% boxprob.Constraints.consLBF = lim_box_free >= 0;

boxprob.Constraints.consBL = box_larger >= 0;
boxprob.Constraints.consDL1 = dim_larger1 >= 0;
boxprob.Constraints.consDL2 = dim_larger2 >= 0;
boxprob.Constraints.consDL3 = dim_larger3 >= 0;

%LP solve
[soln,fval] = solve(boxprob,'Options',options);

```

```

%results

air_per_new = fval/sum(soln.box*box_vol)

for i = 12:210
    % metric-based down-selection
    lol1ol = sum(soln.box)./(1:210);
    [~,I] = maxk(lol1ol,i);

    box_name_d = str2double(box_name(I));
    box_vol_d = box_vol(I);
    box_dim_max1_d = box_dim_max1(I);
    box_dim_max2_d = box_dim_max2(I);
    box_dim_max3_d = box_dim_max3(I);

    box_info = [box_name_d.',box_vol_d,box_dim_max1_d,
                box_dim_max2_d,box_dim_max3_d];
    box_info = sortrows(box_info,[2 3 4 5]);

    box_name_d = box_info(:,1);
    box_vol_d = box_info(:,2);
    box_dim_max1_d = box_info(:,3);
    box_dim_max2_d = box_info(:,4);
    box_dim_max3_d = box_info(:,5);

    %calibrate volume factor for a loose fit
    vol_fac = 1;
    for j = 1:test_order_num

        for k = 1:length(box_name_d)

            pass = false;

            %vet box volume is larger than total order
            volume
            if box_vol_d(k) < order_dim(j)*(vol_fac^3)
                continue
            end

            %vet all SKU(s) can be fitted into the
            selected box
            if box_dim_max1_d(k) < order_dim_max1(j)*
                vol_fac
                continue
            end
        end
    end
end

```



```

        end

        if box_dim_max2_d(k) < order_dim_max2(j)*
            vol_fac
                continue
        end

        if box_dim_max3_d(k) < order_dim_max3(j)*
            vol_fac
                continue
        end

        pass = true;

        %if passing all tests then jump out of the
        loop

        break

    end

    if k == length(box_name_d) && pass == false
        air_per_new_new(j) = 99999;
    else
        air_per_new_new(j) = (box_vol_d(k)-order_dim(j)
            )/box_vol_d(k)*100;
        new_box(j) = box_name_d(k);
        new_box_V(j)= box_vol_d(k);
        new_box_L(j) = box_dim_max1_d(k);
        new_box_W(j) = box_dim_max2_d(k);
        new_box_H(j) = box_dim_max3_d(k);
    end

    end

    flag = 99999;
    out1 = order_info(find(air_per_new_new==flag),:);
    air_per_new_new = air_per_new_new(find(air_per_new_new
        ~=flag));

    air_per_new_avg(i-11) = mean(air_per_new_new);
end
figure
plot(12:210, air_per_new_avg, '*-', 'LineWidth', 1)
xlabel('Size of Carton Set')
ylabel('Air Percentage [%]')

```

```

grid on

pop = [order_info,new_box.', new_box_L.', new_box_W.',
       new_box_H.', new_box_V.'];
[~,idx_ori]=max(soln.box,[],2);

pop1 = [order_info,idx_ori,box_dim_max1(idx_ori),
        box_dim_max2(idx_ori),box_dim_max3(idx_ori),box_vol(
        idx_ori)];

box_data = sum(soln.box);
[idx,C] = kmeans(box_data.',12);

figure
col={'ro', 'r*', 'r+', 'bo', 'b*', 'b+', 'ko', 'k*', 'k+',
     'go', 'g*', 'g+'};
for l = 1:12
    color=rand(1,3,'single');
    plot(find(idx==l),box_data(idx==l),col{l})
    hold on
end
hold on
plot(box_data,'--')
grid on
title('Carton SKU vs. Assigned Order Quantity')
xlabel('Carton SKU')
ylabel('Assigned Order Quantity')
legend('Cluster 5','Cluster 11','Cluster 1','Cluster 2','
       Cluster 3','Cluster 8','Cluster 9','Cluster 7','Cluster
       10','Cluster 6','Cluster 12','Cluster 4')

figure
lol = sum(soln.box);
plot(lol(67:210),'LineWidth', 2)
title('Carton SKU vs. Assigned Order Quantity')
xlabel('Carton SKU')
ylabel('Assigned Order Quantity')

```

Listing A.2: Carton Selection Algorithm Prototype

```

clear all

%box info (pre-sorted)

box_name = ["SFED","J12","MFED","J14","SC","LFED","Sq","LC

```

```

    ", "J18", ...
    "J22", "J64", "MD", "FEDB", "E-C"];

box_vol =
    [440, 750, 2016, 3200, 4928, 1456, 784, 1700, 11907, 20800, 30450, 286, 504, 936]

box_dim_max1 =
    [11, 13, 18, 20, 22, 14, 16, 17, 27, 40, 42, 13, 14, 18].';
box_dim_max2 =
    [8, 10, 14, 16, 16, 13, 7, 10, 21, 26, 29, 11, 12, 12].';
box_dim_max3 = [5, 6, 8, 10, 14, 8, 7, 10, 21, 20, 25, 2, 3, 4].';

box_vol =
    [286, 440, 504, 750, 784, 936, 1456, 1700, 2016, 3200, 4928, 11907, 20800, 30450]

box_dim_max1 = [13, 11, 14, 13, 16, 18, 14, 17, 18, 20,
    22, 27, 40, 42].';
box_dim_max2 = [11, 8, 12, 10, 7, 12, 13, 10, 14, 16,
    16, 21, 26, 29].';
box_dim_max3 = [2, 5, 3, 6, 7, 4, 8, 10, 8, 10,
    14, 21, 20, 25].';

%interface w/ SAP

prompt = 'ord info: ';
ord = input(prompt);

%get ord volume

ord_vol = 0;

for i = 1:length(ord(:,1))
    ord_vol = ord_vol + ord(i,1)*prod(ord(i,2:4));
end

%get ord max dim

ord_dim1 = 0;
ord_dim2 = 0;
ord_dim3 = 0;
temp1 = 0;
temp2 = 0;
temp3 = 0;

temp1 = max(ord(:,2));

```

```

temp2 = max(ord(:,3));
temp3 = max(ord(:,4));

ord_dim1 = max([temp1 temp2 temp3]);

if ord_dim1 == temp1
    ord_dim2 = max([temp2 temp3]);

    if ord_dim2 == temp2
        ord_dim3 = temp3;
    else
        ord_dim2 = temp3;
        ord_dim3 = temp2;
    end

elseif ord_dim1 == temp2
    ord_dim2 = max([temp1 temp3]);

    if ord_dim2 == temp1
        ord_dim3 = temp3;
    else
        ord_dim2 = temp3;
        ord_dim3 = temp1;
    end

elseif ord_dim1 == temp3
    ord_dim2 = max([temp1 temp2]);

    if ord_dim2 == temp1
        ord_dim3 = temp2;
    else
        ord_dim2 = temp2;
        ord_dim3 = temp1;
    end

end

%iteratively select box

%calibrate volume factor for a loose fit
vol_fac = 1.05;

for j = 1:length(box_name)

%    %vet box volume is larger than total order volume

```

```

if box_vol(j) < ord_vol*(vol_fac^3)
    continue
end

%vet all SKU(s) can be fitted into the selected box
if box_dim_max1(j) < ord_dim1*vol_fac
    continue
end

if box_dim_max2(j) < ord_dim2*vol_fac
    continue
end

if box_dim_max3(j) < ord_dim3*vol_fac
    continue
end

%if passing all tests then jump out of the loop
break

%    if box_vol(j) >= ord_vol*(vol_fac^3) && box_dim_max1
(j) >= ord_dim1*vol_fac && box_dim_max2(j) >= ord_dim2*
vol_fac && box_dim_max3(j) >= ord_dim3*vol_fac
%        break
%    end

end

%present result

if j == length(box_name)
    disp('ERROR: Cannot find a suitable box')
else
    disp('Recommended box:')
    box_name(j)
    disp('Air %:')
    (box_vol(j)-ord_vol)/box_vol(j)*100
end

```

This page was intentionally left blank.

Bibliography

- [1] Bin packing problem. https://en.wikipedia.org/wiki/Bin_packing_problem. Accessed: 2019-08-15.
- [2] Bin-packing problem. <http://mathworld.wolfram.com/Bin-PackingProblem.html>. Accessed: 2019-08-15.
- [3] k-means clustering. https://en.wikipedia.org/wiki/K-means_clustering. Accessed: 2019-08-15.
- [4] Packsize: Custom product packaging solutions. <https://www.packsize.com/>. Accessed: 2019-08-15.
- [5] Standard corrugated boxes. <https://www.packagingsupplies.com/collections/corrugated-boxes>. Accessed: 2019-08-15.
- [6] Waters Corporation. Packing procedure. Internal document used for training packers.
- [7] Erick Dube, Leon R Kanavathy, and Phoenix Woodview. Optimizing three-dimensional bin packing through simulation. In *Sixth IASTED International Conference Modelling, Simulation, and Optimization*, 2006.
- [8] A.D. Harlalka. Design and implementation of an rfid-based dock door system at a distribution center. Master's thesis, Massachusetts Institute of Technology, 2018.
- [9] M. Rother and J. Shook. *Learning to See: value-stream mapping to create value and eliminate muda*. Massachusetts: Lean Enterprise Institute, 1999.
- [10] D. Yu. Evaluation of outbound operations improvement projects for distribution centers. Master's thesis, Massachusetts Institute of Technology, 2019.