U.PORTO
FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

Multi-Product Distribution in the Fashion and Retail Industry: A Case Study

Ana Rita Gonçalves Sousa

Dissertation
Master in Modelling, Data Analytics and Decision Support System

Supervised by
Dalila B. M. M. Fontes (a)

(a) University of Porto, School of Economics and Management

2018

Bibliographic Note

Ana Rita Gonçalves Sousa is a Portuguese national born on 28th of February of 1994 in Santo Tirso. In 2012, she joined School of Economics and Management, University of Porto (FEP) and graduated in Economics in 2016. Also in 2016, she enrolled in the Master in Modelling, Data Analytics and Decision Support System at FEP, to which she is now a candidate to obtain the master's degree. During the bachelor, she became a member of a student organization, the Experience Upgrade Program.

In November of 2017, she was admitted as a business intelligence trainee at SONAE 2017 trainee's program. Nowadays, she works as data scientist at the same company.

Acknowledgements

First, I would like to thank my adviser PhD Dalila B. M. M. Fontes for all the guidance, time and patience. Her help was essential throughout the dissertation, without her ideas, high spirits and thoughtfulness this would not be possible.

To my family, specially my parents, I owe the biggest thanks of all because they always supported me in my choices even when they did not understand it, they were always there for me and they did not blame me for my mind absence nor lack of time for them. What's more, they gave me the opportunity to study the things for what I am passionate about.

I would also like to thank all my friends, especially Paulo Pereira, for listening and supporting me, for helping me to keep my mind off of the problems when I need it and understanding my absence.

To the company in this study I want to thank for all the support, help and availability, specially, Pedro Gonçalves, for the help in data collection and all the things he taught me.

Finally, but not least, I would like to thank all people from online forums or websites that started a discussion or answered to it because they really helped me find answers to my questions and were fundamental in my learning process. Even knowing they will not read this, they deserve it.

The support of ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme and the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project POCI-01-0145-FEDER-031821 is acknowledge.

Abstract

This dissertation aims at creating a decision making support tool for downstream managers of a Portuguese fashion company. The case study addresses a multi-product distribution problem considering both product distribution and warehouse purchasing decisions. The distribution problem under study considers a set of different products that is to be distributed from a central warehouse to several geographically distributed stores.

To solve this problem, a mixed integer linear programming model is proposed in which one seeks to find which products are to be sent to which stores in which quantities at minimum cost, while satisfying a large set of constrains such as capacity limits, business rules, and demand among others. The solution approach uses a receding horizon control mechanism that is capable of reducing computational time and enhancing the solution quality. The computational experiments performed show the efficiency and effectiveness of the proposed approach. The solution obtained in the case study used has total annual costs that are 33% lower than those of the company current practice.

Keywords: Distribution Problem; Multi-Product; Optimization; Mathematical Programming, Receding Horizon Control.

Resumo

Esta dissertação tem como objetivo a criação de uma ferramenta que apoie os distribuidores de produto de uma empresa de moda portuguesa na tomada de decisão. O caso de estudo foca um problema de distribuição de diversos produtos onde duas decisões são consideradas, a distribuição dos produtos em si e o nível de compras a realizar pelo armazém. O problema envolve um conjunto de diferentes produtos que têm de ser distribuído, a partir de um armazém central, por várias lojas geograficamente distribuídas. A abordagem escolhida foi a minimização de custos, sendo que todas as decisões tem de respeitar um conjunto de restrições como por exemplo, limites de capacidade, regras de negócio, procura entre outras.

O problema, foi formulado como um modelo de programação linear inteira mista, cuja resolução incorpora um mecanismo de "*receding horizon*". As soluções são obtidas recorrendo ao software CPLEX. As experiencias computacionais realizadas mostram a eficiência e a eficácia da abordagem proposta. Foi obtida uma redução no custo total anual da distribuição dos diversos produtos pelas diferentes lojas de cerca de 33% face, à solução atualmente praticada pela empresa.

Palavras-Chave: Problema de Distribuição, Produtos Diversos, Otimização, Programação linear inteira mista, "*Receding Horizon*".

Index of Contents

Index of Tables

Index of Figures

1. Introduction

The dissertation focus is a specific case study, therefore the data used has been gathered from the company regarding the business year. The company, is one of the biggest Portuguese retail and fashion companies. These Portuguese brand is a pioneer and a specialist on the creation of technical fits. The company operates in the fashion sector, which is a challenging and highly competitive in international terms.

According to the most recent data1 (2015/2016), the retail industry is the biggest industry and second biggest employer in Portugal. This industry is strong and has always been very important to the country, its economy and its development since it employees a lot of people and despite the financial crises of 2007 in which these were severely affected, the sector did not give up and is recovering.

The fashion industry has several specifications, which makes it a highly competitive sector, and in Thomassey (2010) we can read about some of these particularities. The seasonality of sales is one of the most obvious specifications, since it is strongly related to the weather conditions, which are hard to predict accurately and always changing. On the other hand, end-off-season sales, sales promotion and costumer's purchasing power are some examples of the exogenous variables that have great influence in the sector. In addition, as well-known, fashion trends have a major impact in this sector due to two main reasons: (i) it influences the sales directly as most costumers want to buy trendy items and so the company has to keep up with the competition; and (ii) usually the majority of the items is not present in more than one collection – short life cycle – and so there is few historical data to be studied. Finally, the different colours and sizes that need to be available to satisfy the different types of customers lead to a very large number of stoking keeping units (SKU). While, for example, a gas station has about 4 SKU – leaded petrol, unleaded petrol, simple diesel fuel and diesel fuel –, a company like the one in this study can have thousands of SKU

[1] Pordata:
https://www.pordata.pt/Portugal/Empresas+total+e+por+sector+de+actividade+econ%C3%B3mica-2856
https://www.pordata.pt/Portugal/Popula%C3%A7%C3%A3o+empregada+total+e+por+sector+de+actividade+econ%C3%B3mica-3384

since each size of each colour of each product has a different SKU associated. In conclusion, fashion is a very volatile market and the changes are very difficult to predict as they mainly involve people reactions. Thus, in order to be effective when responding to the sector needs, the companies need to have the most adequate tools in order to make the best decisions taking into account the aforementioned issues in the shortest time possible.

This study aims to improve the decision making process regarding the daily multi-product distribution from a central warehouse to several stores in a retail and fashion company. Several companies in different fields face supply chain management problems on their planning routine every day. These problems involve moving raw materials, transforming the raw materials into one or several products types, storing the products, and distributing them amongst the existing stores. Relatively to the part we are studying, the distribution, there are three questions to be answered: when, how, and how much. Some of the studies we will present next, try to answer to more than one of these questions or to answer to questions related with different areas of the supply chain. For example, in vehicle routing problems (VRP) the goal is to answer to the how question, which is to determine the best route for the distribution On the other hand, focusing on the when question, some authors study the best solutions for ordering period. And as a last example, some consider all the questions and all the areas, studying the supply chain management problem as a whole by integrating production, inventory and distribution decisions and determining them together. Nevertheless, our focus is on the when and how much question.

In this Dissertation, we address the proposed problem by first formulating it as a mixed integer programming model[2] that minimizes product distribution related and then solving the model using the software, IBM ILOG CPLEX version 12.7.1. We choose cost minimization over profit maximization even though this is the ultimate goal of all companies because when maximizing profit, we look mostly at the stores that most sell and the product allocation is done based on that, which can compromise some of the company strategies. For example, with the increasingly globalization process, the way companies deal with distribution costs has been gaining more and more importance because these costs have great

[2] A mixed integer programming model is a model in which one or more variables must take integer solution values.

influence on the expansion strategies implementation success. Besides, the distribution costs have a really big impact on the consumer sale price of the product, which means that minimizing them is something that companies seek and prioritize so they can achieve better return margins. With this the company will be able to provide a better service and/or better prices. What's more, by improving margins, it is possible to improve profit, which means that cost minimization it is more versatile in global terms.

Motivation

This study is the last step of the Master in Modelling, Data Analytics and Decision Support Systems and the choice of theme came from the will of learning more about decision support systems and optimization. In the daily life of companies there are a lot of decisions that need to be made, some are easy and others are really hard. What makes most of them so hard is the complex and conflicting information that has influence on the decisions, even if at first sight it seems easy, there is a lot behind it. Nowadays, in most of the companies, even really important decisions are made based on rules of thumbs or using Excel or some similar basic software, thus taking into account only some of the important characteristics. Hence, only by chance the companies may be making the best decisions, so there is a lot to be improved. The reasons for that may be related to the lack of access to it and/or to the fact that the available tools are hard for the managers to understand and use. As a result, creating a customized resolution process, simple enough to be used by the managers, is a great incentive.

The choice of multi-product allocation problem was born of the contact with a retail and fashion company and as a suggestion of Professor Dalila B. M. M. Fontes. This problem, in the company it is being addressed, has a simple resolution process based on some rules that are applied daily. The fact that we are studying a real life problem faced by several companies in different sectors is a big incentive since we hope to provide the company with a tool that is capable of improving their decision making process and thus obtaining better results.

More specifically, my motivation relies on the following presented reasons. As was mentioned before, the impact of distribution related costs on sale price is large and that represents a big weight, particularly on a competitive sector as the retail and fashion. Thus, trying to decrease the costs is also a great reason for studying the proposed problem. In addition, the cost minimization is also very important because of the growth of globalization. In addition and also previously mentioned, the importance of distribution related costs increases when products are sent to other countries. In this situation if the quantities sent are not suitable, correcting it will be even more expensive because, for example, it might imply another shipment or it might take the stores to incur in unnecessary costs. Thus, the growth of globalization ends up being a strong motivation for studding the distribution problem as well.

Outline of the Dissertation

The reminder of the dissertation is organized as follows. In Chapter 2, we discuss the most important references to our work and introduce some important concepts and similar problems. Chapter 3, provides a detailed description of the problem. The data collection, the mathematical model, and the proposed to solve the model are presented in Chapter 4. Chapter 5 describes the computational experiments and discusses the results obtained. Finally, in Chapter 6 some conclusions are drawn and future directions are pointed out.

2. Literature Review

Many aspects of supply chain management (SCM) problems have been studied and reported in the literature. There are several studies on the SCM sub-problems involving different point of views and different approaches. Regarding the product allocation problem there are several versions since it depends on the type of allocation that is to be made. For example, Chen et al. (2014) studied the production allocation of several products to different production facilities on the apparel industry. On the other hand, Luo et al. (2017) explored product allocation to shelfs, which is related to space availability and organization of the stores and warehouses.

The oldest study on distribution problems we found is due to Stephen C. Allen *a*nd dates back to 1958, (Allen, 1958). In this pioneer work, the author solves a stock redistribution problem by minimizing the costs incurred with the redistribution between the several locations.

Accordingly to Cretú and Fontes (2017), the problem considered in this dissertation, is an inventory routing problem that ignores the routes. Thus, the closest approximation is when only direct shipments are considered because each retailer has its own shipment so the routes have much less influence. In addition, Galleo and Simchi-Levi (1990) proved that when the quantity shipped to each retailer is close to a full load, the direct shipment is efficient.

These type of problems have been largely studied and so there are several variants, considering several features. Some of the details that lead to so many different problems are, for example: planning period (single, multiple or infinite), the demand type (deterministic or stochastic), the route type (direct or with transhipment), the decision making process (centralized or decentralized) and the frequency of shipment (single or multiple). Next, we define the context of the problem being address regarding these features.

2.1 Problem Context

Starting with the determination of products distribution we have: static allocation and dynamic allocation. Kumar et al. (1995) say that static allocation is determined based on inventory levels for all customers simultaneously, before the vehicle of transportation leaves the warehouse. On the other hand, dynamic allocation is determined sequentially, based on inventory levels at the arrival of the transportation vehicle to each customer. In their paper, they study static and dynamic policies for replenishing and allocating products to several customers on a fixed delivery route and they assume that each customer faces independent, normally distributed period demands. As our approach is going to be to solve the problem for all stores at once, we have static allocation.

Also related to the allocation, we have the shipment frequency, (Speranza and Ukovich, 1994). Thus, if a product is always shipped with the same time interval it is said to be single frequency and if it is shipped with different time intervals it is said to be multiple frequency. Another distinction made by the authors is whether the products shipped together have different shipment frequencies or not. If products with different frequencies can be shipped together, then there is time consolidation; otherwise there is frequency consolidation. In this study, the products have multiple frequency and can be shipped together, thus there is time consolidation.

A definition for the two types of existing demands can be found in Xu et al. (2007): can be: deterministic –where the predictions are assumed to be the true values and so inventory levels are known and stochastic – where the predictions are not assumed to be the values and so inventory levels are unknown. These authors explored an optimal product distribution under the vendor managed inventory problem. In such problem, at each customer the vendor faces a trade-off between replenishing the current customer at a high level and reserving more product for the following customers that may or may not bring more profit, in other words, dynamic allocation process. In addition, according to Kang and King (2010), demand can also be dynamic and thus, change from a period to another or static and thus, remain the same. In the specific case study we address, the demand is deterministic and dynamic.

Regarding the route type, in Eliiyi et al. (2011) we find the transhipment concept. Transhipment happens when the products are shipped to an intermediate location (transhipment depots) and then from there shipped to the retailers. This is used as a mechanism for correcting differences between demand and inventory levels, or for example, when there is the need of changing the means of transportation. According to the authors, the goal is to determine the replenishment quantities of the retailers and product quantities to be sent to the transhipment depots at each shipment such that the cost is minimized. In our case, the routes are ignored because the products transportation is outsourced and only the origin and destination need to be specified. Thus, the route does not have influence in the transportation cost and direct shipment is the closest approximation.

One last important issue regarding problem context is the decision-making environment that can be centralized or decentralized, accordingly to Lee and Jeong (2010). In centralized decision-making environment, there is only one decision-maker responsible for solving and defining the solution for the warehouse and retailers, the objective is global. On the other hand, when it is decentralized, the warehouse and the retailers determine their solutions independently in order to achieve their local objectives. Usually, a centralized decision making process has better results regarding cost minimization than a decentralized one, mainly because of competition drives the decision makers to not share information about their own business. In our problem the decision-making environment is centralized.

2.2 Methodology Context

According to Bradley et al. (1977), mathematical programming is the best developed and most used technique of scientific decision making management approaches. Mathematical programming is very used to solve optimum allocation problems while having limited resources in a competitive environment, which means while satisfying a set of constraints that describe the addressed problem nature. In addition, when only linear functions are used in the problem formulation, we have, as in our case, a linear programming model. With mathematical programming we can help managers to better understand the consequences of their decisions. If we take the time to know the problems nature and evaluate the quantitative method role, is possible to improve the decision-making quality. There are three decision-making types, strategic, tactical and operational. Decision as resources allocation throughout a range time horizon fall into the tactical type and are one of the mathematical programming specialities. Our case study mathematical programming formulation is presented in section 4.2.

Despite the above mentioned advantages, mathematical programming has a disadvantage, it assumes certainty and in the present case study we have a factor of uncertainty, the predicted demand. Thus, in order to treat the uncertainty factor, we are going to combine two complementary disciplines, mathematical programming and control theory. In Dauod et al. (2017), it is possible to find the definition of receding horizon control (RHC), one of the most important concepts we have. The RHC technique divides the original problem into smaller time frames. The idea is to, in the first step, solve the fixed horizon optimisation problem in order to get a set of solutions but consider only the solutions related with the current time instance. Then time advances and at each time instance the process is repeated like a moving window.

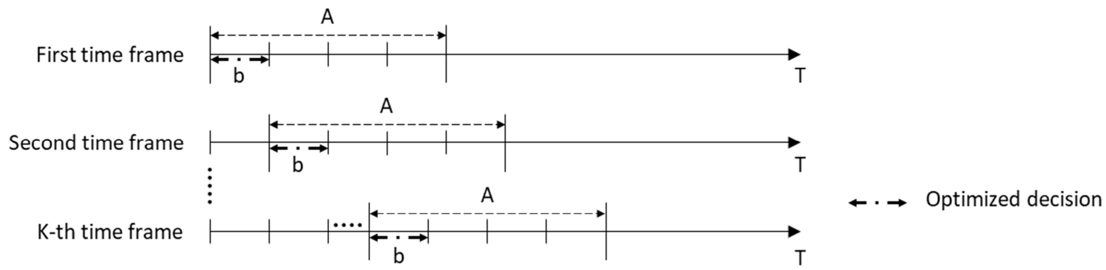

Figure 1 Receding Horizon.

As Figure 1 shows, although the problem is solved for “A”, the optimization horizon, only “b”, the optimal decisions implementation horizon, is implemented. Considering “T” as optimization horizon instead of a midterm period would be extremely time consuming and eventually it would use long term demand previsions without updates, which is too risky. On the other hand, solving the problem only for “b” is also risky because it does not take into account the future and the decisions would be optimized locally. With RHC technique is possible to find, for all time instances, solutions that consider not just all the optimization horizon but also updated information. Besides being very good to solve real time control problems, RHC also helps to improve the computational demand and enhances the solution’s quality. In this case study, RHC is used to update the stocks level and thus, as mentioned before, mitigate the effects of the uncertainty brought by the predicted demand.

2.3 Similar Problems

In this section, we are going to present similar problems with similar features and different contexts, with different features and similar context, and with similar features and similar context. Different problem are going to be discussed to show different sectors where product allocation is very important and that each one has its own particularities.

Of the first works, one of the most studied product allocation type, Bassok and Ernest (1995). This authors consider a multi-product and space allocation problem in the Soft Drink Industry, where the distributor has limited transportation capacity with a fixed known sequence of costumers; in other words a previously determined route. However, their demand is unknown, this information is only available when the distributor arrives at the customers, which means that when arriving at the first costumer its demand becomes known but not the demand of the following costumers. The first decision that needs to be made by the distributor is how many units to allocate to a certain customer given that the demand of the next customer(s) is unknown. This decision is made taking into account that the product on the truck might not be enough to satisfy the demand of the costumer(s) still to be visited and if there is not, there are penalties. The second decision is the space allocation, however, as the authors treat the two decisions separately, using different approaches we are only going to discuss the first part which is most relevant. The first part of the problem is solved by a dynamic programming approach, based on Brumelle and McGill (1993), in order to get expected profit maximization. In Brumelle and McGill (1993), for determining the airplane seat allocation that maximizes revenue, a stochastic dynamic programming is used. Their problem is solved recursively. On the other hand, Bertazzi et al. (2005) studied the same problem, but with known inventory levels, in other words, deterministic demand. They solved their problem through a constructive heuristic that at each iteration the supply for one retailer is inserted in the solution. First, they decompose their main problem into smaller ones and then the sub problems are solved hierarchically. After having a solution they try to improve it by iteratively changing the solution for two retailers. If no better solution is found they return to the original solution.

As an example of static allocation, Topaloglu (2005) studied the allocation of product to different regional markets together with its production allocation. However and as expected, we will focus on the part of the problem that is similar to ours, the product allocation to the markets. The product in this work can only be stored in the plants because the product is perishable and so the regional markets do not have the necessary conditions to store it properly. This restriction leads to a trade-off between what proportion to save in the plants and what proportion sent to the regional markets because it is important to ensure that there is enough product to supply a profitable market in future periods. Although this restriction represents a big difference between Topaloglu (2005) and our problem, the implications of it are small. The most important and significant similarities between the problem addressed by Topaloglu (2005) and our problem are described below. When planning the delivery of the product to customers, the author takes several things into account such as: inventory levels, forecasts of customers' demands and production capacity. Despite the fact that we do not have production capacity, we do have storage capacity which can be thought of more or less in the same way. The author formulates the problem as a dynamic programming model and then solves it by using a linear approximation of the objective function. As the inventory levels and the set of time periods are independent, the initial inventory levels available to be sent to the markets in a certain period of time are used as state variable. The original functional form chosen for the value function was a concave function, which means that an incremental unit of product stored decreases the marginal profit. However, solving the value function optimality's equation with classical backward recursion techniques is, accordingly to the authors, usually not the best option because of the "curse of dimensionality" (the data becomes sparse with increase in dimensionality). On the other hand when using the objective function approximation, the solution comes down to solving sequences of small dimension cost network flow problems in order to maximize the profits. The solution method starts with a set of value functions approximation and at each iteration it is tried to improve the approximations by updating the value function, similarly to what is done with RHC, with solutions determined using random quantities of product to produce at each plant.

Kang and King (2010) considered a supply chain, with one supplier and a group of retailers, in two levels and determined the quantity and timing of product to be delivered to the customers, with a static allocation process. First a solution for the quantity of product to be delivered is determined and then its timing. For the first phase, which is of most interest for us, their objective was to minimize the sum of inventory holding and handling costs assuming that: (i) the demand is dynamic and known, (ii) the transportation vehicles have limited capacity and can visit more than one customer in the same route, and (iii) the routes are already determined so they are ignored. The transportation costs are only considered in the process' second phase. Besides the fact that in our problem we focus on the quantity of product to send to each retailer and the quantity of products that the warehouse should order while Kang and King (2010) focused on the quantity of product to send to each retailer and the timing for this delivery, we have the same or similar assumptions. To determine the quantity of product to be delivered to the customers, Kang and King (2010) used simple and well tested algorithms, known for achieving very good solutions on dynamic lot sizing problems and all of them are applied based on a heuristic approach. In addition, besides the two level approach, the problem is first solved ignoring the capacity restrictions and then considering it. Regarding optimal solution algorithms, two different approaches were used, Wagner and Whitin (1958) when ignoring capacity restrictions and Lambert and Luss (1982) when considering it. On the Heuristics side, Silver and Meal (1973) approach is used both times, first on its original form and then modified so it could consider the vehicles capacity limits. Wagner and Whitin (1958) algorithm core logic relies on the following steps: first the algorithms look for the possible ways of demand satisfaction at each moment, then they calculate the cost for each one of the distribution's policy, after that, the distribution's policy with lower costs is chosen and the process is repeated for the next time instance. On the other hand, Lambert and Luss (1982) have a more developed approach. They narrow the search for possible policies solution to the area close to the extreme point of the feasible region. Then the original problem is divided into smaller ones in order to explore extreme values for the decisions variables and compare objective function values. The solution is constructed step by step, it starts with a solution for the first considered period and, at each exploration, adds solutions for the other periods according to the best objective function values. As mentioned before, on Heuristic's side, Silver and Meal (1973) approach is used. This approach is based on the determination of the average cost per period. While the average cost of the period □ is higher than the average cost of period □ $-$ 1, the process continues; otherwise the process stops and it starts again from period □ forward. On the modified Silver and Meal (1973), the logic remains the same but the final solution is altered in order to respect the capacity limits. Thus, when the delivery quantities obtained, by the original Silver and Meal (1973), for a period exceed the vehicle capacity, the delivery quantity is reduced and the not urgent deliveries are delayed.

Contrarily to the previous examples, which involve concepts that are mostly related with mathematical and dynamic programing approaches, receding horizon control (RHC) can be applied to all optimization approaches that consider multi-periods. For example, Goodwin et al. (2006), apply RHC to a mine planning exploration. The authors define a set of depths as the mine states on specific mine locations, and then the evolution of the mine state is given by a dynamic model that uses mining actions as an input. These input is in its turn given by RHC which is used to decide, under some constrains, which is the mining action to happen in a certain moment. In addition, Fortenbacher et al. (2018) used RHC for deciding strategies for distributed battery storage. Their goal is to maximize the photovoltaic utilization and minimize battery degradation while respecting all constraints. Considering long horizons it is not a feasible option since predictions dispatch methods are only available for short horizons due to details like the volatility of weather prediction. Thus, the problem is divided into a master problem, storage planning, and then in several sequentially solved sub-problems, which reflect the operational strategy; and using RHC they transfer the previous optimization cycle results to the consecutive one as an initial state input. To conclude, with a very similar approach to the one in this case study, Dauod et al. (2017), consider a pharmacy replenishment planning problem, where they have two decision variables: (i) the reorder level of dispenser inventory and (ii) the number of canisters to use. In order to solve their problem, the authors use a mixed integer programming model in order to minimize the total replenishment costs and apply RHC technique to reduce the computational demand and enhance the solution's quality. RHC mechanism allows the solution's quality enhancement because while traditional planning models cannot detect dynamic changes in real-time inputs, RHC can. In the problem formulation, two index are considered, dispensers and time, since that for avoiding drug contamination, each dispenser and canister, can only contain one type of medication and so, contrarily to our case study, there is no need for the type of medication to be other index. Besides this, they consider capacity limits, operational costs and assume that there is sufficient inventory in storage during all the replenishment process. In addition, Dauod et al. (2017), use CPLEX to solve their mixed integer programming model.

3. Problem Description

This dissertation addresses the single-warehouse multi-retailer multi-product distribution problem, Figure 2. The warehouse is where the process of distribution begins as it is where all products are kept since the production is finished until it is decided in which quantity and to which store they go to. As previously mentioned in Chapter 2, we face a static allocation meaning that we determine the allocation for all stores at the same time and a deterministic dynamic demand meaning the demand is known although it changes over time. Thus, the problem main goal is to determine which products are sent to which store and the respective quantities in order to minimize all the distribution related costs.

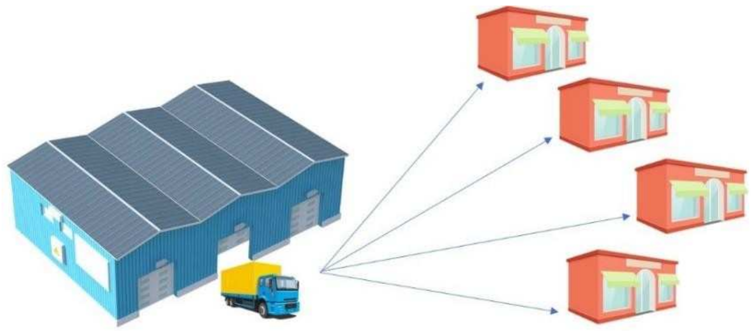

Figure 2 Distribution Schema.

3.1 Case Study Description

As previously mentioned in Chapter 1, a company like the one in this study with thousands of SKUs since each size of each colour of each product is a different SKU. Products are divided along several dimensions, as illustrated in Figure 3. The first and more general division is the category: bottom parts, top parts, and accessories. Then it follows the product families: (i) jeans, (ii) jackets and coats, (iii) sweaters and cardigans, (iv) t-shirts and polos, (v) shirts, (vi) skirts and dresses, (vii) shoes, and (viii) accessories. And so on, until the last and most detailed level division, the SKU, which is the product type division level considered by the company when allocating products to stores. They have a system where they group the stores by type and needs. The stores groups are created according to each family type, which means that a store can be part of a group regarding jeans and be part of other group regarding accessories. Each group defines, for its stores, what are the references the stores must have.

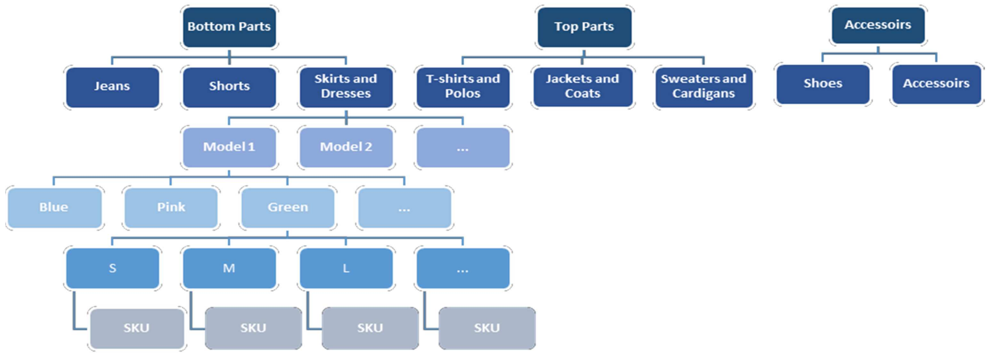

Figure 3 Product Division.

Currently, the company decides how to allocate the products on a daily bases. However, business rules require two weeks' worth of product to be kept at all times in each store. Besides this, the company has a product's diversity policy, which means that at all time periods each store must have available products of every family. In addition, there are some time periods in which some products must be sent to every store. This happens because collections are not sent to the stores all at once. For each collection, Spring/Summer and Autumn/Winter, the new products arrival to the stores are divided in six moments and at each one of these moments every store receives new products even if they do not need them. Yet, some stores receive more quantity of products than others due to both their dimension and potential sales. Henceforth, there is a guideline for the quantity of new products that must be sent to each store at each new product entry. These business rules represent some of the problem constrains and the solution must respect them.

Equally important and also representing a set of constrains are capacity limits, regarding both product storage and expected demand. For each product type, and each store, the quantity stored must satisfy upper and lower limits. These reference limits differ, and usually do, amongst stores and product families. In addition, there is limit on the maximum total quantity of inventory that a store can bear. Besides the above mentioned limits, only part of the inventories at the warehouse can be allocated to the stores since the stock at the warehouse is also used to satisfy other demand channels. Naturally, warehouse inventory is also limited by its storage capacity. Moreover, the company's current distribution contract has no limit neither for the number of boxes to be transported nor for the number of transports to be made. Hence, besides when it is planned, the products can be shipped when their storage levels are unexpectedly low at the stores. In addition, after product expedition, it is the distributor's responsibly to manage the delivery.

There are, however, two other transportation issues, namely: box content and lead times. For transportation, the box content is usually mixed, in other words, all products types can be shipped together. This has to do with the way in which the piking (separation of products) is made. There is picking by store and picking by line, however both, mix in the same box all type of products to be shipped to each store. The great majority of the boxes used has the same size and the mean number of products that a box can take is 20. However, it varies depending on the product. For example, for winter coats, the worst case scenario, this number comes down to five, while for t-shirts it goes up to 100. The only and rare exception to the previously mentioned and usually used box is when there are a lot of accessories to be sent to a store because, in that case, they prefer to ship it all together in a smaller box. Nonetheless, its transportation cost is the same as that of the larger box. The lead time is one day for most of the stores considered, in other words, the products that are shipped in one day arrive at the stores in the following day. The exceptions, some stores in Spain, have two days lead time. Thus, the products that are shipped in one day arrive at the stores two days later.

Thus, based on all the rules, capacity limits, and the fact that there is detailed information about the sales plan and demand forecasts for each store, product distribution managers decide the main product allocation. First, product distribution managers define an automatic provisioning for a relatively long period of time taking into account the expected demand. Then, every day, at the end of the day, when real sales become known, stock levels are updated. With this new data, every day, the product managers check if the planed allocation is working as expected and make the changes they see fit. These adjustments are important because the actual sales may and probably will, differ from what was expected. For example, whenever a store sells more of one product than forecasted, the store needs extra units of that product in order to comply with the two weeks demand rule. On the other hand, whenever a store sells less of one product than forecasted, the store inventory level is above the required one. In the second case, when the store is close to its capacity limit, the store has to return at least part of it to the warehouse.

3.2 Detailed Problem Description

In this section, we provide the detailed description of the specific problem addressed in this work. We are going to consider a representative approximation of the real problem since considering all the complexities associated with the original problem would require information that it is not available and resources that we do not have, such as computational power. Despite that, and according to the company product distribution manager, we do have a good representation of the real problem, which allows us to infer good quality conclusions that can be extrapolated to the larger frame since all the main features are considered. This detailed description serves also as an introduction to the mathematical notation used in the mathematical programming model that is discussed in the next chapter.

To begin with, there are three key aspects that have great influence on the problem since almost all of the problem's features depend on them. These aspects are: the stores to consider, the product type division and time (horizon and period).

In this case study, we have one central warehouse and 72 geographically dispersed stores in Portugal and Spain. Although there are many more stores in these and other countries, we consider only stores that have the same business rules and use the same transportation means. Also, Portugal and Spain represent the two most important markets for the company, which provide us an excellent business sample.

As mentioned before, the company considers SKU as product division level. However, they do that because of fashion, style, and store type requirements. For example, some stores sell more small sizes while other sell more large sizes and some sell more products of dark denim while other sell more of light denim. This type of features require a sensibility that it is hard to pass to the model and so we think it is better to let that type of decisions to specialists, while providing them with the information on the quantities to be sent to the stores at each time instance. Hence, we choose the division level of product families instead, since we consider it represents well the company's products diversity offer.

Regarding the time period, we choose one business day since in the company decisions are made daily. The planning horizon was chosen to be one year, so that the seasons are considered, however we are left with 250 days, after removing weekends and holidays.

Therefore, decisions need to be made regarding the quantity of each of the eight products to send to each of the 72 stores in each time period of the planning horizon.

Three features are going to be considered in a slightly different way than what they are in the original problem. First, we will only consider the part of the warehouse referent to the stores that are being studied in this work. As previously mentioned, only part of inventory levels at the warehouse can be allocated to the stores since the stocks kept at the warehouse are also used to satisfy other channels of demand. Thus, the initial inventory levels considered at the warehouse refer to the initial inventory levels available to send to the considered stores. The inventory levels at each period are given by the model. The warehouse storage capacity limits will also be adjusted to the one available for the considered stores. Secondly, the smaller transportation boxes are not considered because their use is extremely rare and the transportation cost is the same as that of the other boxes. Thirdly and lastly, we will not take into account the delivery lead times. As we consider that two weeks' worth of product must be kept at all times and all stores, the probability of selling two weeks of expected demand in one or two days is extremely low and so their impact on the problem is not significant. Besides this three features, all the other rules and capacity limits are considered as described in the previous section.

In addition, we are going to consider other two new features, that is, features that product managers do not take into account. One of these features is related with one of our goals, which is to minimize the operational costs by including them in the decision making process. Therefore, the costs that we propose to be taken into account are: storage costs at the warehouse and stores, product handling costs at the warehouse and stores, and transportation costs. For the warehouse storage and product handling costs there are reference values since the company has already studied them. The handling cost are measured by the relationship between the hourly income of employees and the quantity of time that an employee spends handling a box or a product unit. This cost differs between the warehouse and stores. On the one hand, at the warehouse, the loading task is automated, which decreases the handling time and therefore, the handling costs. On the other hand, the employees have considerably more tasks associated with product handling. The piking task requires lots of time since it includes several sub-processes to ensure that the picking error is minimized. In contrast, at the stores the product handling tasks is manual and includes receiving, properly storing and/or displaying the products, which implies an easier handling process. Regarding storage costs, they differ not just between warehouse and stores, but also amongst the stores. Despite the fact that the warehouse storage costs is known, the company did not studied yet the storage cost at the stores. Thus, the squared meter price of each store's rent is going to be used as a proxy of this cost. And last but not the least, the transportation cost, which is a per box cost, even if in one day there is only one box to be shipped the cost remains the same. Transportation costs do not depend on travelled distance nor on truck capacity occupied and there are no minimum or maximum number of boxes requirements for transportation to take place.

The need for the other new feature we propose, appeared with the data exploration and first computational experiences performed. After the data exploration, it was clear that the warehouse kept higher levels of stock than predicted and stores ended up having considerably more stock than what they were supposed to. Hence, we decided to add a new decision variable to the model in order to define, optimally, what products should be bought by the warehouse and in which quantities.

3.3 Contributions

This work major contributions are twofold. On the one hand, we include costs in the decision making process. On the other hand, we also optimize buying decision.

The company present decision making process does not have into account the several costs involved. The decision on the quantities of product to send to the stores is based, besides business rules, only on historical data and stock levels. Costs management is a major issue on the success level of any company and the current decision making process is neglecting it. This overlooked fact represents the first improvement opportunity. Thus, the decision making process that we propose includes, as described in the previous sections, amongst other important variables and constrains, the product distribution related costs.

The second improvement opportunity, is the incorporation product's buying decisions since, as already discussed, we have observed that, usually the stock levels kept at the warehouse and the stores are much higher than it needs to be. Sometimes, the stores end up having enough stock to satisfy their demand for two months which represents four times what they should have. This situation may be explained, at least in part, by the fact that product allocation and product buying decisions are made by different teams, and thus, the inventory levels might be reflecting lack of information on the product's buying decisions. Therefore, we will propose to simultaneously decide: (i) what products are sent to the stores and in which quantities; and (ii) what products should be bought by the warehouse and in which quantities.

4. Methodology

4.1 Data Collection and Processing

Data collection and processing represent a really important task in this work. Thankfully, the company had available all the required data so there was no need for artificial data generation. However, most of the collected data was not in the format that we needed. Data processing, was performed using software R (R Core Team, 2014).

The first step was to verify if there were all types of data available for all stores, products and periods. As we will discuss, with detail, in the next section, we consider input data like: initial inventory level, predicted sales, effective sales and capacity limits. Unfortunately, there were some stores that did not have all the input data available. This is due to the fact that some of these stores were opened during the analysed year. Consequently, as these stores were on a probation period, their data was not consolidated, which could skew the results. Thus, we decided to remove these stores from set of stores to be considered. Stores that were franchised part of the year and then bought by the company have also been removed, since they operated part of the year on a different regime and under different rules. Besides this two situations, there is a third one, in which some stores had incoherent data and so were also removed from the considered set. Therefore, of the 72 previously mentioned stores, we consider 51.

After collecting and analysing the data for the 51 stores, some processing was need. For example, stocks level had to be aggregated since instead of SKUs we consider product families. In addition, as mentioned in Chapter 3 we consider 250 eligible days for distribution so we had to identify and remove data regarding weekends and holidays, regarding both predicted and effective sales. This transformation was made in order to make comparison possible and to update stocks at the end of each time period. As a way to reduce computation time we complied, for each day, the total demand of the following fourteen days. Since we must satisfy the business rule that imposes a constraint inventory level for each product and each store capable of satisfying the predicted demand for the next fourteen days, this way saves computational time.

Besides these transformations, we had to format the data in the way that CPLEX requires. There are three types of inputs as illustrated in Figure 4: (i) a one-dimensional input, in which each row represents a product; (ii) a two-dimensional input, in which, as before, each row represents a product and each column represents a store; and (iii) a three-dimensional input that in addition to products and stores, also considers time (three columns).

1 dimension

Products	Values
1	Values
2	Values
3	Values
...	Values

2 dimensions

	Stores	Stores	Stores	Stores
Products	1	2	3	...
1	Values	Values	Values	Values
2	Values	Values	Values	Values
3	Values	Values	Values	Values
...	Values	Values	Values	Values

3 dimensions

Products	Stores	Time	Values
1	1	1	Values
1	1	2	Values
1	2	1	Values
1	2	2	Values
...	...	...	Values

Figure 4 Input Data Types.

To conclude, the last transformation was the data anonymization, we use numbers to refer to family products and stores. These numbers were assigned randomly.

Regarding data storage, we use Microsoft Excel and each time instance has an Excel file with the data inputs relative to it. CPLEX reads the inputs from the Excel files and then writes the results in them as well. Part of the data inputs are constant, for example capacity limitations, while others change over time, for example predicted demand. Therefore, using R we were able to create automatically the 250 different excel files with the constant inputs and then, according to each time frame, write in the respective excel file the other inputs.

4.2 Mathematical Model

In this section, we describe the mathematical programming model developed for solving the multi-product distribution problem being address. Table 1, summarizes the notation used.

Table 1 Mathematical notation.

Sets and Indices	Description
T	Set of time periods (days) in the planning time horizon, indexed by □ ∈ ▪ ;

I	Set of product families, here and hereafter referred as products, with |□| = □, indexed by □ ∈ ▪ ;
J	Set of stores, with |□| = □ and □′ = □ ∪{▪ } representing the set of stores together with the warehouse, from now on referred as facilities, index by ▪ ∈ ▪ ′;
Parameters	**Description**
□□□□	Initial inventory level of product □ ∈ ▪ at facility ▪ ∈ ▪ ′;
□□□□	Predicted sales, in units, of product □ ∈ ▪ at store ▪ ∈ ▪ in time period ▪ ∈ ▪ ;
□□□□	Minimum required quantity, in units, of product □ ∈ ▪ , that must be sent to each store ▪ ∈ ▪ in time period ▪ ∈ ▪ ;
□□□□	Effective sales of product □ ∈ ▪ at store ▪ ∈ ▪ in time period ▪ ∈ ▪ ;
□□	Quantity, in units, of product □ ∈ ▪ that can be transported in single box;
□□□□	Minimum number of units of product □ ∈ ▪ that facility ▪ ∈ ▪ ′ must have;
□□□□	Maximum number of units of product □ ∈ ▪ that facility ▪ ∈ ▪ ′ can have;
□□	Maximum number of boxes that can be transported to store □ ∈ ▪ ;
□□	Minimum number of boxes that need to be transported to store □ ∈ ▪ ;
□□□	Minimum quantity of stock, in units, in each facility □ ∈ ▪ ′;
□□□	Storage capacity, in units, of each facility □ ∈ ▪ ′;
□□□	Storage cost per box at store □ ∈ ▪ ;
□□(□+□)	Storage cost per unit at the warehouse;
□□	Sum of the warehouse and stores handling cost, per unit;
□□	Transportation cost per box;
Decision Variables	**Description**
□□□□	Quantity, in units, of product □ ∈ ▪ sent to store ▪ ∈ ▪ in time period ▪ ∈ ▪ ;
□□□	Quantity, in units, of product □ ∈ ▪ purchased in time period ▪ ∈ ▪ ;
□□□□	Quantity, in units, of product □ ∈ ▪ returned to warehouse by store ▪ ∈ ▪ in time period ▪ ∈ ▪ ;

Auxiliary Variables	Description
□□□□	Stock, in units, of product □ ∈· at facility · ∈·′ at the end of time period · ∈·;
□□□	Number of boxes sent to store □ ∈· in time period · ∈·;
□□□□	Existing number of boxes of product □ ∈· in store · ∈· at time period · ∈·;
□□□	Number of boxes returned (to the warehouse) by store □ ∈· in time period · ∈·;

Next the mathematical formulation is presented, first the objective function and then the constraints.

Objective Function:

$$□□□.□ = □□ \sum_{□∈·} \sum_{·∈·} \sum_{·∈·} (□_{□□□} + □_{□□□}) + \sum_{·∈·} □□_{□} \sum_{·∈·} \sum_{·∈·} □_{□□□}$$
$$+ ·· \sum_{·∈·} \sum_{·∈·} · + ·· \sum_{·∈·} \sum_{·∈·} (· + ·) \qquad (1)$$

Subject to:

$$\sum_{□∈·} □_{□□□} \geq □□_{□} \qquad ∀·∈·′, ·∈· \qquad (2)$$

$$\sum_{□∈·} □_{□□□} \leq □□_{□} \qquad ∀·∈·′, ·∈· \qquad (3)$$

$$□_{□□} \leq □_{□} \qquad ∀·∈·, ·∈· \qquad (4)$$

$$□_{□□} \geq □_{□} \qquad ∀·∈·, ·∈· \qquad (5)$$

$$□_{□□□} \geq □□_{□□} \qquad ∀·∈·, ·∈·′, ·∈· \qquad (6)$$

$$□_{□□□} \leq □□_{□□} \qquad ∀·∈·, ·∈·′, ·∈· \qquad (7)$$

$$□_{□□□} = □_{□□(□-1)} + □_{□□□} - □_{□□□} - □_{□□□} \qquad ∀·∈·, ·∈·, ·∈· \qquad (8)$$

$$□_{□(□+1)□} = □_{□(□+1)(□-1)} + □_{□□} - \sum_{□∈·} □_{□□□} + \sum_{·∈·} · \ldots \qquad ∀·∈·, ·∈· \qquad (9)$$

$$\sum_{\square \in \cdot} \square_{\square\square\square} \leq \square_{\square(\square+1)(\square-1)} \qquad \forall \cdot \in \cdot, \cdot \in \cdot \qquad (10)$$

$$\square_{\square\square\square} \geq \square_{\square\square\square} \qquad \forall \cdot \in \cdot, \cdot \in \cdot, \cdot \in \cdot \qquad (11)$$

$$\square_{\square\square} \geq \sum_{\square \in \cdot} \frac{\square_{\square\square\square}}{\square_{\cdot}} \qquad \forall \cdot \in \cdot, \cdot \in \cdot \qquad (12)$$

$$\square_{\square\square} \geq \sum_{\square \in \cdot} \frac{\square_{\square\square\square}}{\square_{\cdot}} \qquad \forall \cdot \in \cdot, \cdot \in \cdot \qquad (13)$$

$$\square_{\square\square\square} \geq \frac{\square_{\square\square\square}}{\square_{\square}} \qquad \forall \cdot \in \cdot, \cdot \in \cdot', \cdot \in \cdot \qquad (14)$$

$$\square_{\square\square\square} \geq \sum_{\square+1}^{\square+14} \square_{\square\square\square} \qquad \forall \cdot \in \cdot, \cdot \in \cdot \qquad (15)$$

$$\square_{\square\square}, \square_{\square\square}, \square_{\square\square\square} \geq 0 \; \square\square\square \; \square\square\square\square\square\square\square \qquad \forall \cdot \in \cdot, \cdot \in \cdot \qquad (16)$$

$$\square_{\square\square\square}, \square_{\square\square\square}, \square_{\square\square\square} \geq 0 \qquad \forall \cdot \in \cdot, \cdot \in \cdot' \qquad (17)$$

The model and respective constrains were generally thought so that it would be possible to use and adapt it to similar problems. Thus, due to the fact that, for example, in our problem there is no transportation limits, the corresponding constraints will not be implemented in CPLEX (IBM ILOG, 2009).

Equation (1) describes the cost minimization nature of the problem. The costs considered are: (i) handling cost, which is a per unit cost in both warehouse and stores; (ii) storage cost, which for the warehouse is per unit of product and for the stores is by square meter occupied; and (iii) transportation cost which is fixed per box.

The first six constrains refer to the capacity restrictions. Inequalities (2) and (3) ensure that, in each period, the inventory for both the stores and warehouse are within their respective upper and lower limits. Similarly, inequalities (4) and (5) ensure that the transportation limits are respected. They, respectively, refer to the maximum and minimum quantity of boxes that can be sent to each store. Inequalities (6) and (7) ensure that the inventory held at each facility for each product in each time period is within the specified limits.

Constrains (8) and (9) are the balance equations for the warehouse and stores, respectively. The maximum quantity of product that can be sent to all stores is given by

Inequalities (10), while Inequalities (11) ensure that the quantity of each product sent to each store in each time period is at least the minimum required quantity. In addition, the use of the correct number of boxes in the transportation is ensured by inequalities (12) and (13) for sending and returning products, respectively, and for inventory storage by Inequalities (14). Inequalities (15) ensure that each store in each time period, keeps as inventory at least the inventory level of each product required to satisfy the next predicted demand, for the next fourteen days (business rule). Lastly, constrains (16) and (17) define the variables nature. Note that, as the inventory is obtained by adding and subtracting integer values, the variables associated with it will always be assumed integer values as well.

As mentioned in Chapter 3, decisions are made daily, thus the proposed model is to be solved for every day of the planning horizon. Therefore, inventory levels at the stores at the beginning of every day are given as an input, by updating the inventory levels at the end of the previous day obtained by the model. As a result, at the beginning of each day, the stock levels have to be updated using the real sales rather than the expected demand: $\square_{\square\square\square} = \square_{\square\square\square} + \square_{\square\square(\square-1)} - \square_{\square\square(\square-1)}$. The warehouse stock levels do not need to be updated since we defined its purchases as a decision variable and so the stock levels resulting from previous decisions are the real ones.

4.3 Solution Approach

Theoretical background

As mentioned before and according to several authors, see for example Dauod et al. (2017), mathematical programming is widely used to solve optimization problems like ours. Besides this fact, we want to find an optimal solution by creating a solution process that faithfully represents the real problem. Thus, for these reasons, the chosen approach for this case study was to use mathematical programming combined with receding horizon control.

We choose to use a commercial software, such as CPLEX (IBM ILOG, 2009) to solve the mixed integer linear programming model proposed in Chapter 3. This software is one of the best known and most widely used, since it is efficient and robust in solving such models.

The CPLEX approach to the models is a little bit like a black box, however we know that to solve mixed integer programming problems like ours it uses either branch and cut search or dynamic search.

The branch and cut algorithm is a combinatorial optimization method that consists in the following elements: linear programming relaxation, branching, cuts, and heuristics. In addition, CPLEX provides parameters to enable, disable and/or control these elements.

In a branch and cut algorithm (Elf et al., 2001), the original problem is divided in a series of continuous subproblems. The set of these subproblems can be represented by an enumeration tree, where each subproblem is called branch and cut node, from now on referred as node. The enumeration tree initializes with a root node that represents the continuous relaxation of the original problem, and then, the exploration of the nodes starts. There are four types of nodes. The node that is currently being explored is called current branch and cut node. The active nodes that still have to be processed, and the already processed nodes that can be fathomed or not fathomed. A problem is fathomed when: the local lower bound (lpval) of his subproblem is equal or greater than the global upper bound (gub), or the subproblem becomes infeasible, or the subproblem has been solved, being its solution a feasible solution of the original problem. The nodes are generated in a branch step that usually appears when the bounds of one variable are modified. For example, when exploring the lower and upper bound of the parent node limit, which means one child gets the value of the lower bound and the other gets the upper bound value. On the other hand, the cuts role is to limit the search space by eliminating part of it, by adding a new constraint to the model. Cuts, typically, prevent fractional solutions, which reduces the number of branches needed to solve the problem. No admissible solution to the original problem is ruled out by a cut. At each node the solution formed can be of three types: all-integer solution, infeasible solution or another fractional solution. In the first two cases the node is fathomed. In a fractional-valued integer variable presence, the algorithm decides when is best to branch and when is best to cut. If possible it decides to cut because it reduces the problem size. The process has to be repeated until one of the first two solution is reached. There are two major keys aspects in branch and cutting: the computation of a global upper bound (gub) and a local lower bound (lpval). Ideally, an incumbent solution – the best known solution to satisfy all the integer requirements – is quickly found so that it can be used by the algorithm as global upper bound when trying to find better solutions. The gub value is then compared with the global lower values (glb), which is given by the minimum lower bound of all the active and current nodes. When the glb is lower than the gub, a better solution has been found and the glb solution becomes the new gub solution, which means becoming the new incumbent solution. When the list of nodes to solve is empty, the gap[1] – the reflex of the progress towards finding optimality – will be zero and the optimality of the incumbent solution will be proved. Despites, often CPLEX terminates the solution process when the gap has been brought to 0.01% or less because usually much computation is required in order to prove optimality.

However, sometimes, the algorithm cannot find a solution, which was our case. We left the software running for an hour and no incumbent solution was found, Figure[2] 5. Thus, it became unfeasible to solve the problem in this way, so the next step was to try dynamic search.

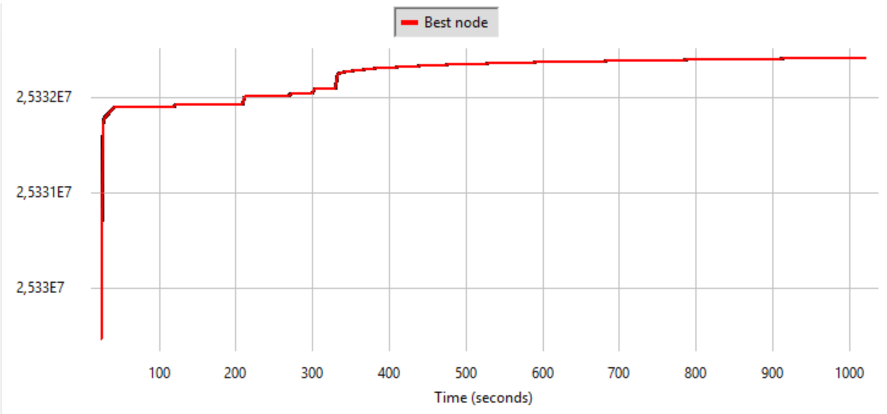

Figure 5 Solution process representation of CPLEX when using traditional branch & cut method with no heuristics.

The dynamic search algorithm is based on the same concepts and logic as branch and cuts but it has a different implementation. The advantage in this method is the fact that it achieves solutions faster than the conventional branch and cut method. Still, this was not

---

1 |□□□□□□□□□□−□□□□□□□□□□□□|
1□−10+|□□□□□□□□□□□
□|

2 For the purpose of allowing a better comparison, all the presented graphs consider the same time interval.

good enough since the solution found was far from the optimal solution, when CPLEX stopped the process the gap was at 97%. As it can be seen in Figure 6, the best integer line is far away from the best node line showing that the difference is considerably big.

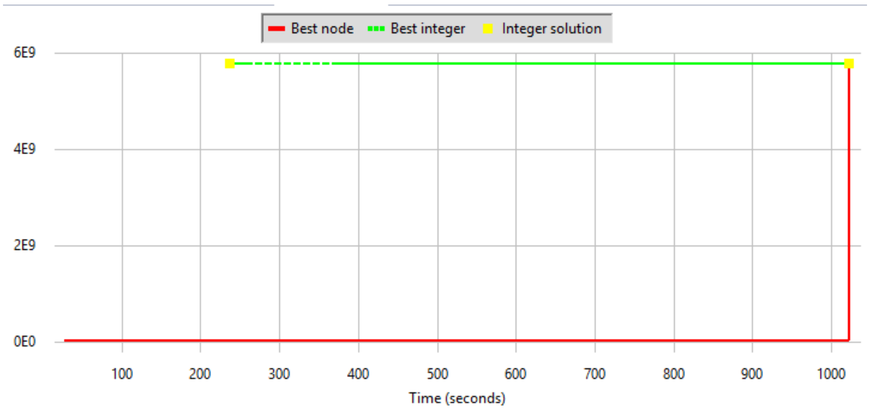

Figure 6 Resolution process representation of CPLEX when using dynamic search method with no heuristics.

Besides dynamic search, there is one last thing that can help improving our solution. When CPLEX cannot find a solution for a node there are two options: branching or heuristics[3]. Thus, in order to find the best possible solution, we allowed CPLEX to use heuristics whenever the algorithm was having difficulties in finding a solution for a specific node. Although it does not replace branching, heuristic can quickly and inexpensively find a good approximate solutions to a subproblem. In addition, a solution found in this way is treated as any other feasible solution. Thus, CPLEX integrates heuristics in the branch and cut method, which allows to speed the final proof of optimality. Allowing this, truly improved our solution. It enabled us to get really close to the optimal solution, the final gap is always under 0.05%. As it can be seen in Figure 7, the best integer line overlay the best node line showing that the difference is quite small.

[3] According to Shah and Oppenheimer (2008), heuristics are techniques that reduce and simplify the effort associated with a certain task. This techniques can quickly find satisfactory solutions, however, without theoretical guarantees of optimality.

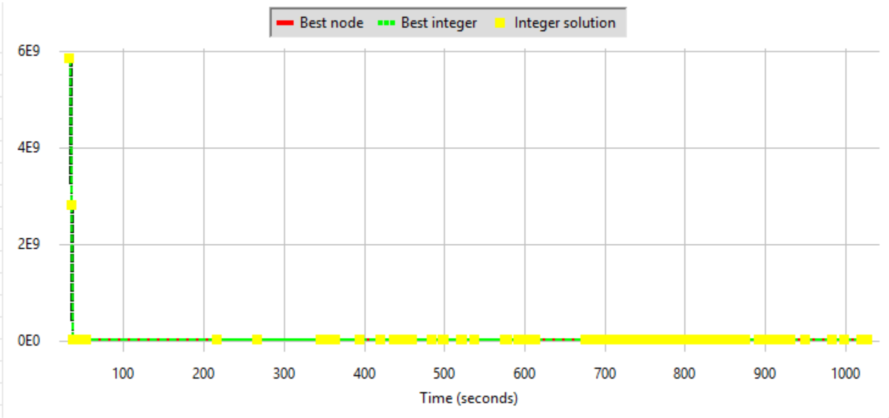

Figure 7 Solution process representation of CPLEX when using dynamic search method with heuristics.

Hence, we decided to solve the problem using dynamic search and heuristics to help finding nodes solutions when there was no better way. Besides this, there are two more important CPLEX parameters to discuss. Regarding the trade-off between feasibility and optimality, we choose to balance them, rather than emphasizing one or the other because it makes the algorithm look for fast proof of optimality without forgetting to take effort in finding high quality feasible solution. To conclude, we allowed the use of presolver and/or aggregator[4] more than one time so that a good initial linear relaxation is achieved in the algorithm first building block where the initial problem is simplified.

Adjustments and Framework

After discussing the theoretical part of implementation, we will now discuss a limitation that had to be imposed to solution process in order to make it feasible and will briefly discuss the constructed framework.

Initially, our idea was to solve the problem without any external limitation or relaxation however, with the first experiments we realised that it would not be feasible given the time required. We tried to solve the problem without a time limitation and although CPLEX gets close to an optimal solution quickly, it takes a long time to ensure optimality. Usually, it is really difficult to improve a solution when it already has a gap close to or of 0.01% and so CPLEX finishes the process but, it still stays on that level too long before

[4] Presolver and aggregator are the two available pre-processing methods on CPLEX and they try to reduce the coefficients of constraints.

stopping to try to get a better solution. The computational time that CPLEX takes to solve a problem does not depend only on the size of the model, but also on things like the gap between the incumbent and optimal solution, and the proportion between basic and fractional solutions. Hence, different data sets mean different computational times, which in turn means each day can require different computational times. Thus, we decided to limit the computational time of CPLEX to 1000 seconds. With this limitation we look for, in the available time, the best possible solution. As reported in section 5.2, the final optimally gap is always below 0.05%. When the limitation being imposed to the model is regarding the final gap, the recommendation value is of 0.05%, therefore, from the assigned solution when CPLEX reaches the computational time limit, the possible improvement is not much.

To conclude the implementation discussion we briefly describe the framework of the solution process, which mainly consists on a cycle. Everything starts with the main script that controls the cycle. First, the main script calls the model file, which is where we have the parameters and variables definition, the objective function, and the constraints. Then, the main script reads from the main data file which is the name of the instructions data file to consider at that time instance and calls it. The instructions data file tells to the model script from which excel file, page and range the input data is going to be read. After this, the model script starts the solution process for that time instance. When the final solution is found the results for the decision variables are written on different pages of the same excel file from where the input data was read. Besides this, it is also written, in the excel file of the next time instance, the stock level of the current time instance for each product at each facility and with this information the real stock levels are instantaneously updated since the remaining required information for the update is already in the file (section 4.1.). The process described is repeated until the last time instance is considered.

The solution process is completely automated. Thus, after defining for which time instances we will solve the allocation problem, and, consequently, having the corresponding inputs on the excel files, in order to start the resolution process, we only have to call a command in PowerShell. When the process is finished, the solutions for the several decision variables will be saved on the excel files and so we can analyse the results.

5. Computational Experiments and Results

5.1 Algorithm Test

To ensure that the chosen approach was correct and that CPLEX was solving the problem as expected, a small example of the original problem was solved using other software. This is very important because it allows for results comparison.

One of the simpler ways to solve small linear programming problems is using the "Solver add-in" tool of Microsoft Excel. This tool has available three possible resolution methods: Generalized Reduced Gradient Nonlinear, Evolutionary, and Simplex LP. The first two methods are able to solve linear and nonlinear problems but the solution they find might not be optimal. These algorithms compromise optimality in exchange for speed in the resolution process. On the other hand, Simplex LP only allows linear functions. However, it assures that the solution obtained is a globally optimum one. Thus, as our problem is linear and we want to find optimal solutions, we choose the Simplex LP method to solve a small representative example of our original problem.

For this test the considered problem included the warehouse, two stores, two products and two time instances. The test was performed for two different examples and, in both, the solutions obtained by CPLEX and by Microsoft Excel were exactly the same, Figure 8.

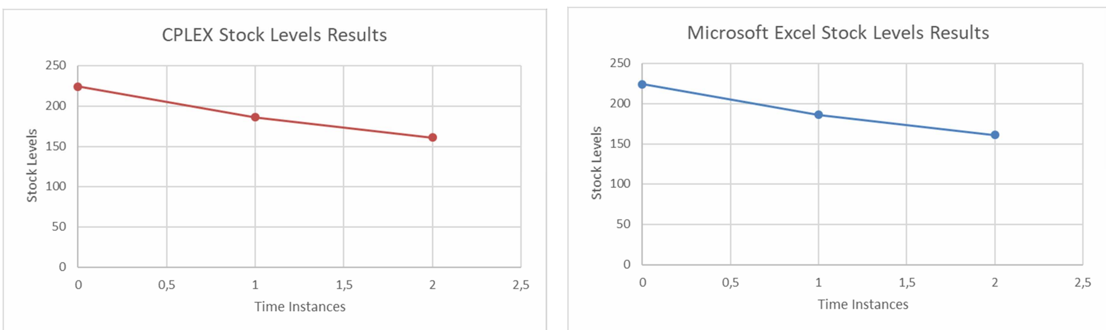

Figure 8 Test CPLEX and Microsoft Excel results example.

5.2 Results Analysis

Computational experiments were performed not only to validate our model but also to analyse its performance and usefulness for the company. Thus, in this section, we will discuss the most important results. The result analysis was performed using software R (R Core Team 2014).

First Experiment

The first analysis is a comparison between the decisions and associated costs obtained by solving our model, and those of the company. As mentioned before, the instances solved here involve deciding on a daily basis, over a time period of one year: (i) the quantities of eight different products to be sent from a central warehouse to 51 stores (retailers) and (ii) the quantities of the same eight products to be bought by the warehouse. The analysis considers 240 days, since only for these the next fourteen days predicted demand are available. Hence that, there is a business rule that forces inventories levels at stores to be at least equal to the next fourteen days predicted demand. Thus, only the first 240 days can be analysed.

The first results presented are warehouse inventory levels and store average inventory levels. In Figures 9 they are reported as a percentage of the facility capacity, for both warehouse and stores, respectively.

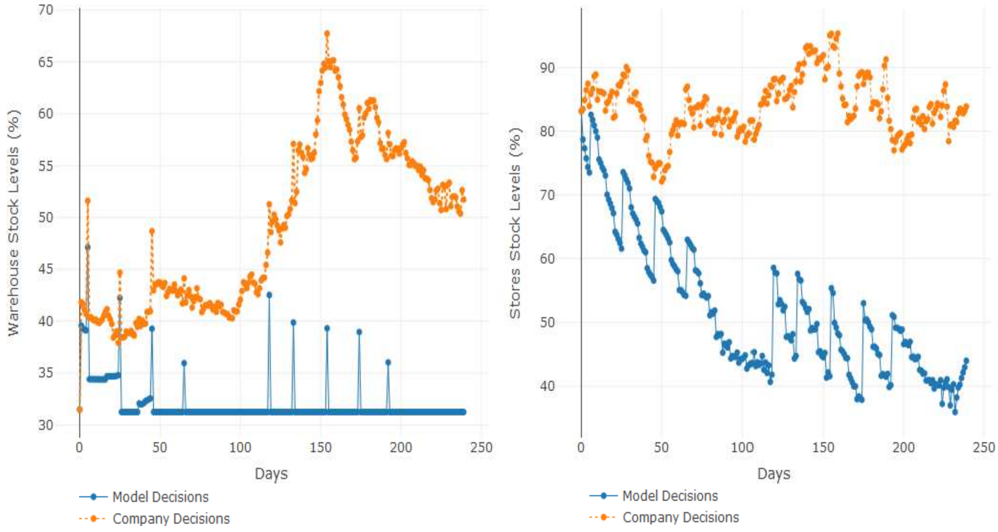

Figure 9 Warehouse and Stores Stock Levels.

As it can be seen, excluding some specific cases, the model decisions for warehouse inventory levels are steady, typically, being at the required minimum level. The nine spikes on the model decisions refer to moments where, there is a minimum required quantity of products to be sent to the stores, which are the entries of new products to the stores and so, consequently, the warehouse has to buy and store those new products before they are sent to the stores. On the other hand, warehouse inventory levels kept by the company were much higher then what they needed to be throughout most of the year. In addition, an exponentially increase can be observed during the first seven months of the year. Then, there is a break around summer promotions phase and after this, it starts growing again for a small period of time until it begins a decreasing tendency at the end of the year. The exponentially growth of stock levels throughout most of the year is something that it is expected by the company and it is something they are trying to change. The moments were the stocks levels were lower, are also expected since they correspond to summer, Christmas and winter promotions.

Contrarily to what happens with the warehouse inventory levels, the company average inventory levels of the stores are steadier than the ones derived from our model decisions. However, similarly to what happens with the warehouse inventory levels, the company stores average inventory levels are much higher than what they needed to be. The new products entry at stores spikes are also very visible in the company store average inventory of the model. In addition, the model average inventory levels of the stores show a descendant tendency which, is a reflex of a more controlled environment at the warehouse and of the synchronization between sent quantities and stores needs according to predicted demand.

In order to better explore the obtained results, we are going to use something similar to ABC analysis, which is an inventory categorization technique (Benito and Whybark, 1986). In our case stores will be categorized based on the mean value of the shopping cart. Our division is going to result into comparisons between the stores with a shopping cart with higher average value, group A, versus the ones with a shopping cart with lower average value, group B, and between the products with shopping cart with higher average value, and the products with shopping cart with lower average value.

The different stores type are presented in Figure 10, each graph shows data regarding a representative store of each type. The left hand side graph shows group A, the group of stores with a shopping cart with higher average value, which means these type of stores achieve higher values of shopping cart medium value selling fewer products. While the right hand side shows group B, the group of stores with a shopping cart with lower average value, which means these type of stores sell the less, both in value and in quantity. This last group is, usually, the reflex of recent openings, which are situation where the stores are trying to gain space in a new market and, consequently, sell less. Similarly to what happened with the previous graphs, the stock levels percentage is relative to the maximum capacity of each group.

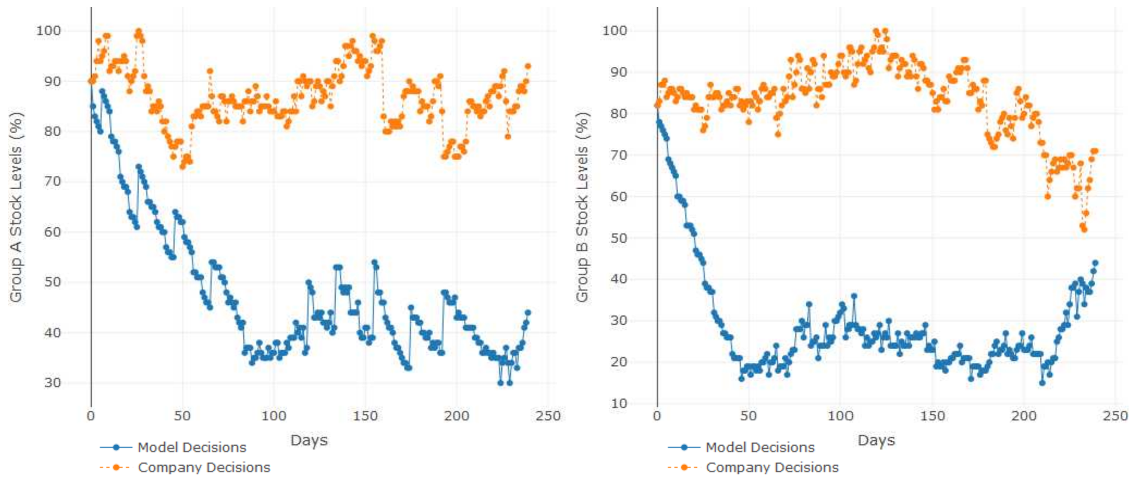

Figure 10 Group A and B Stock Levels.

As it can be seen, group A has, at least for most of the year, the same tendencies for the company and model decisions. The biggest discrepancy between company and model decisions happens in the first months since the descendent tendency was present for considerably more time in model decisions than on company decision. This situation happens because the initial stocks are an input and so are much higher than what they needed to be. Besides this and excluding the fact that it appears that the model anticipates the spikes a little, the ups and down happened more or less at the same moments. This is related, at least in part, with moments where there is a minimum required quantity of products to be sent to the stores and promotions, as was discussed in the previous analysis. On the other hand, in group B the tendencies of company and model decisions are different. In the beginning of the year, while the company decisions lead to steady stock levels, the model decisions lead to an exponential decreasing tendency, which as mentioned before, has to do with the initial stock. After that, the model decisions lead to steadier stock levels while company decisions start a tenuous growing tendency followed by a descendent tendency. Finally at the end of the year, while the company is still on a descendent tendency, the model decisions lead to a growing period. This tendency is probably related with the phase of the year, which is Christmas, which in turns is followed by winter promotions. The effect of this phase is not as visible in some stores as it is in others because some stores have less selling potential and so require less preparation for this year phase when compared with stores with best-selling potential. Nonetheless, the most important and visible aspect is common to both graphs. As expected, this analyse reveals that the company keeps much more stock than needed. As a result of this, an exponentially descendent tendency is always found in model decisions in the beginning of the year.

In addition, while through company decisions, the stores daily keep, on average, stock levels of 84% of its maximum capacity, through model decisions, the stores daily keep, on average, stock levels of 52% of its maximum capacity, Table 2.

Table 2 Summary of Model and Company Stock Levels.

	Min.	**1st Quartile**	**2nd Quartile**	**Mean**	**3rd Quartile**	**Max**
Model Decisions	36%	44%	49%	52%	58%	83%
Company Decisions	72%	81%	84%	84%	87%	95%

We believe that the main reasons for the company stores stock accumulation are related with: first, a previous stocks accumulation at the warehouse and second, the decision of how many references a stores should have. The previous stocks accumulation at the warehouse is related with two things: desynchronized information in buying decisions and difficulties in draining out-season products. The decision of how many references a stores should have, has influence as well because thinking in that, the distribution product managers end up allocating more product to the stores than what was initially defined according to the store's needs. Therefore, consequently, they end up overlooking the business rules that requires two weeks' worth of product to be kept at all times in each store, which leads to a higher stores stocks accumulation than needed.

Product type comparison is presented in Figure 11. Besides the two different product types, the graphs also show their behaviour in different stores type. We used the same groups used in the previous analysis and while product 6 represents one of the core products, product 7 is a niche type product and so it represents a smaller part of sales. As in the previous graphs, the stock levels and effective sales appear as a percentage of the maximum capacity of each store for products 6 and 7, respectively.

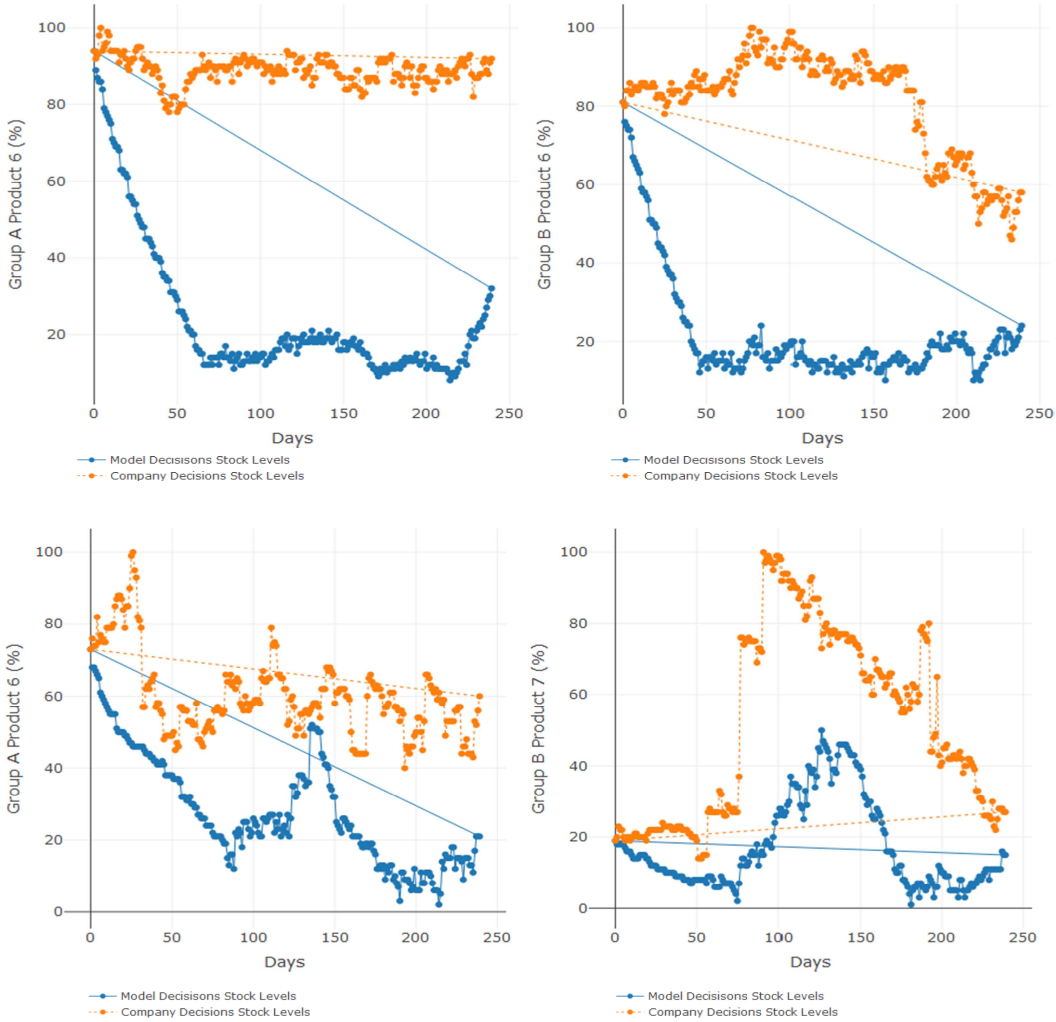

Figure 11 Groups A and B, products 6 and 7 Stock Levels.

Regarding product stocks levels behaviour, it seems to be similar to what was already discussed regarding the overall stores stock levels. For most of the year, the company and model tendencies are more or less the same, the model inventory levels being lower than the company inventory levels. However, it is possible to see that product 6 and 7 behave differently. Product 7 varies through the year, having spikes, while product 6 is steadier after the first decreasing tendency.

With the model decision, in 97920 (height different products × 51 different store × 240 days) decision moments, there are 39 moments of shortages. The first shortage is an isolated case and there is no explanation for it. In that day in that specific store and product, the effective sales were 10 times higher than what was predicted. This phenomenon was not a shortage for the company because as said before they keep much more stock than what they need to. All the other shortages happened in time period 232, which was a campaign period. The impact of this campaign in the sales is hard to predict since is a very recent type campaign in Portugal and so there is little data to study and predict on. Thus, as the campaign was very successful the sales exceeded the predicted demand in several product types and stores. Once again, this was not a problem for the company because they have an excess of stocks at the stores and they deliberately reinforced the stocks where it was necessary. The reinforcement of the stocks during campaigns can be considered by the model if the company wants it. However, the company might not want to sell more than a certain quantity of product in a certain campaign since it is selling with less margin and so less profit. Therefore, we consider that these situations are not problematic. The model allocation is good and shows that stock levels can be much lower and still satisfy the demand.

The costs analysis is presented in Figure 12. The left hand side graph provides the total costs for both model and company decisions; while the right hand side graph shows the ratio of these costs. In our model formulation, we did not take into account a shortage cost because the company did not study it yet and so there is no information about it. However, as discussed above, the two situations where there was shortage, are very specific and would have a low impact in the model costs. Furthermore, as it is very clear, the model costs are lower than the company costs throughout most of the year, which means we were able to reduce, considerably, the distribution operational costs as expected. Overall, we were able to reduce the costs in about 33%. Therefore, it is safe to say that even when considering a shortage cost, the model total costs would be lower than those of the company.

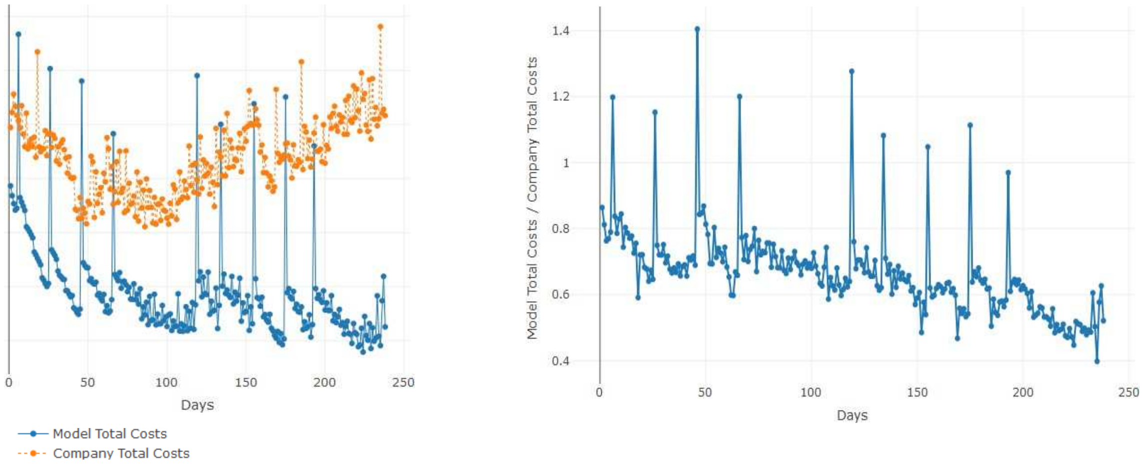

Figure 12 Total Costs analysis Experiment 1.

In the left side graph, it is possible to see that both situation show ups and downs over the year and, as before, the spikes in the model costs are related with the moments where there is a minimum required quantity of products to be sent to the stores. Additionally, in the company total costs, there is a growing tendency after the first year trimester, while in the model total costs tend to stay in the same range after the initially decreasing tendency. The right side graph reinforces that idea since the costs reduction is higher at the end of the year. Nonetheless, there are moments where the model costs are higher than the company costs. These moments are the moments where there are minimum required quantity of products to be sent to the stores.

The summary of the total costs ratio is presented in Table 3. As the total year reduction, the mean cost reduction achieved was of 33% and, at least, in three quarters of the year the reduction achieved was of 28%.

Table 3 Summary of Total Costs Experiment 1.

	Min.	**1st Quartile**	**2nd Quartile**	**Mean**	**3rd Quartile**
Cost Reduction (%)	60	40	34	33	28

To conclude the first experiment analysis, in Table 4 shows the summary of CPLEX time performance and solutions quality. For most of the days CPLEX stopped the optimization process when reaching the time limit. Nevertheless, the solution final gap was always below 0.05%, ensuring almost optimality.

Table 4 Summary of CPLEX Time Performance and Solutions Final Gap, Experiment 1.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max
Time Performance	38.84	1001	1000	814.22	1002	1009
Solutions Final Gap	0.005%	0.010%	0.017%	0.019%	0.025%	0.049%

Second Experiment

As lead time was not considered when deciding which products and in which quantities should be bought by the warehouse, a second experiment, where the real warehouse stock levels were considered as inputs, was performed. By considering the real warehouse stock levels as inputs, we are considering lead times since when deciding the warehouse stock levels, the company took into account the lead times. Also, this allows to observe the influence of higher warehouse stock levels in the store replenishment decisions.

As can be seen in Figure 13, the difference between the model inventory levels of the stores in the first experiment and the second experiment is small, and the tendencies are exactly the same. With this we can infer that not considering lead times has a small impact in stores inventory levels. This was expectable since warehouse storage costs are considerably lower than those of sores and so the model keeps the maximum stock possible at the warehouse.

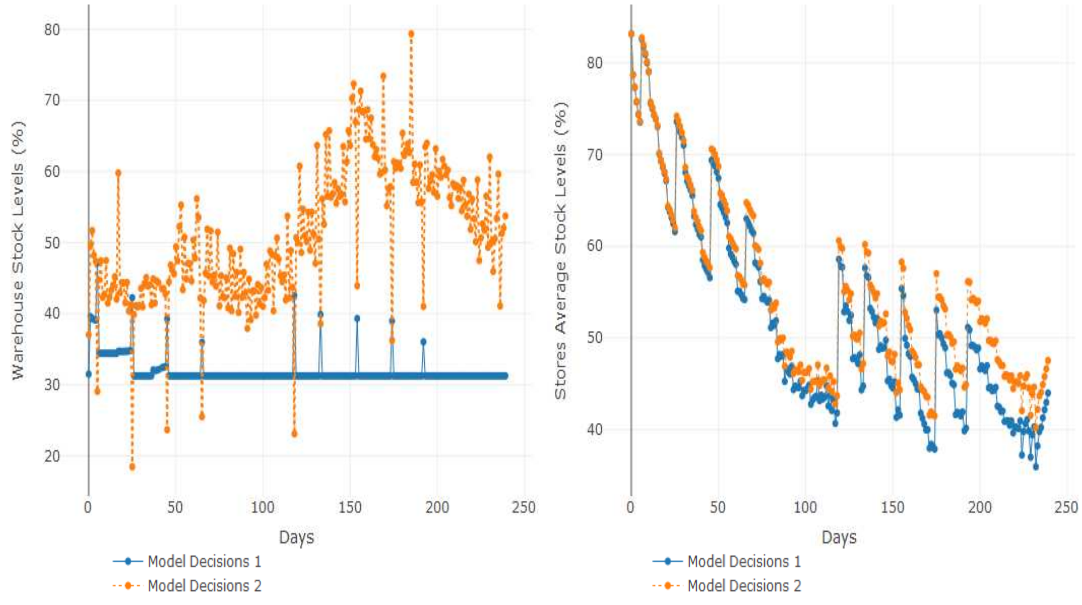

Figure 13 Warehouse and Stores Stock Levels.

Figure 13 also shows the inventory levels at the warehouse for experiments 1 and 2. As it can be seen the inventory levels in model 2 are much higher than the ones in model 1, similarly to what happened in the first experiment. This difference explains the larger total costs of experiment 2 and the incensement towards the end of the year when comparing to experiment 1 (see Figure 14).

The cost analysis also confirms that although the stocks levels at the warehouse are much higher in the second experiment, its influence in the store's inventory levels and costs is small, see Figure 14. In the left hand side graph it is possible to see that the tendencies are the same as the ones observed in Figure 12. On the other hand, the right hand side graph presents the relation between the model experiment 2 total costs and company total costs, and between the model experiment 2 total costs and model experiment 1 total costs. These two ratios are presented in Table 5 as well.

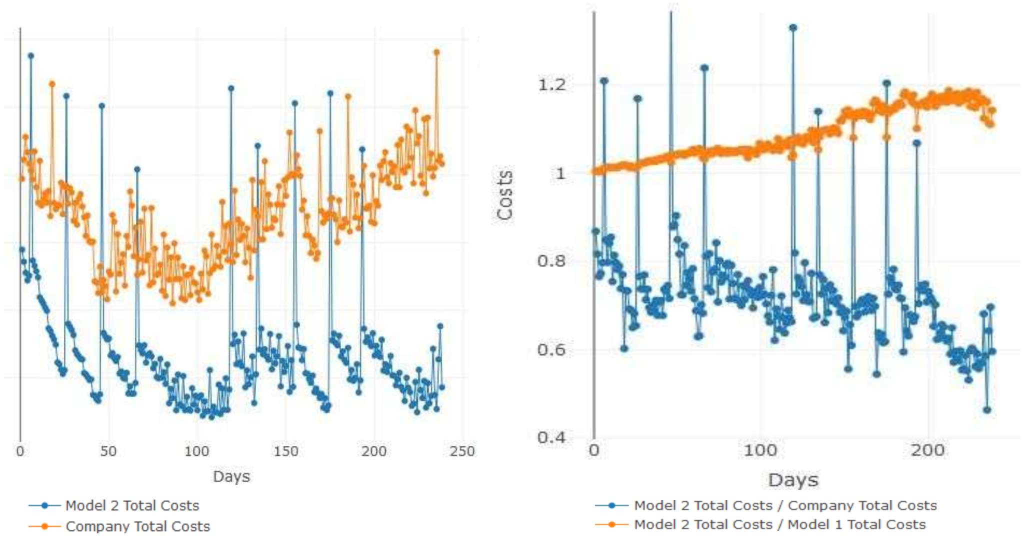

Figure 14 Total Costs analysis Experiment 2.

In the overall of the year, with the second experience, we were able to reduce the total costs in 28%, which means that even not considering purchasing decisions at the warehouse, we were able to reduce considerably the total costs. Note that disregarding purchasing decisions implies larger inventories at the warehouse, which implies an increase in total costs. In comparison with the first experiment the total costs increase, on average, 8% and as it is possible to see in Figure 14. The impact of the inventory levels at the warehouse is higher at the end of the year. The growing tendency, when comparing the two experiment costs, is expectable since we previous observed a higher difference in the stores average inventory levels of the two experience also at the end of the year.

Table 5 Summary of Total Costs Experiment 2.

	Min.	1st Quartile	2nd Quartile	Mean	3rd Qu.
Cost Reduction between Model 2 and Company (%)	54	33	29	28	25
Cost Ratio between Models 2 and 1 (%)	1	1.04	1.07	1.08	1.15

With the second experience model decision, in 97920 (height different products × 51 different store × 240 days) decision moments, there are 35 moments of shortages. Even thought the growth at the stores stock level is small, this growth prevented four moments of shortage and decreases the number of units in which shortages occurs.

Similarly to the first experience, to conclude this analysis, in Table 6 a summary of CPLEX time performance and solutions quality is shown. Similarly to the first experiment, for most of the days, CPLEX stopped the optimization process when reaching the time limit. However, the average time and first quartile where lower. This can be explained by the fact that the problem and this model, no longer include purchasing decisions. Despite and also similarly to what happened before, the solution final gap was always below 0.05%, ensuring almost optimality.

Table 6 Summary of CPLEX Time Performance and Solutions Final Gap, Experiment 2.

	Min.	1st Quartile	2nd Quartile	Mean	3rd Quartile	Max
Time Performance	2	194.9	1002.2	730.8	1003.4	1011.2
Solutions Final Gap	0.0006%	0.009%	0.015%	0.016%	0.020%	0.041%

6. Conclusion

This dissertation addresses a specific multiproduct distribution problem in the fashion and retail industry that is closely related to direct shipping inventory routing problems. With the purpose of improving the company's current process of allocation decision making, we formulated the problem as a mixed integer linear programming model. In order to take into account all the significant variables for the problem definition a detailed research was performed and even variables that, currently, are not taken into account by the company, were included. To solve the proposed model we used CPLEX.

In the beginning of this work, we expected to be able to cut down the costs and, eventually, to point out improvement points or to propose changes to current policies and strategies. Thus, the obtained results have been compared with those of the company current practice. During this comparison we observed that, as expected, the company keeps much more stock than needed in both warehouse and stores. This situation is something that the company already knows and is trying to correct. The implications of the excessive stock levels are vast and have the long term consequences. With excessive stock levels, the company incurs not only in unnecessary costs in the present (mainly due to storage costs), but also in future costs as products became out-seasoned, typically, implying lower margins sales. Besides this, with excessive stock levels the company needs more space for storage at the stores, which requires larger storage space and thus higher costs. All in all, it is very important to have well-structured buying policies because they have a lot of indirect effects and so their impact is much higher than the one seen at the first sight.

As the stock levels excess was expected, in order to find out the ideal warehouse stock levels, we included the purchasing decision in our model as a decision variable. As a result, we were able to prove that with much less sock it was possible to satisfy the demand. On average, we achieve a reduction of 32% in inventory levels at the stores, and in warehouse, 17%. Thus, we were able to reduce the inventory levels and still satisfy the demand.

Regarding the costs, we observed an improvement of 33% in total costs. However, in about 3% of the days, the obtained solution implied higher costs averaging 19%.

Nonetheless, this situation ends up having a small influence since in the remaining 97% days of the year, the obtained solutions lead to average costs savings of 35%.

The verified shortages represented 0.043% of the total sales. However, as mentioned before, this happened during an extremely successful campaign that could be taken into account by the model.

We also conducted an experiment in which purchasing decisions were considered as inputs. We were able to conclude that the impacts of having larger inventories at the warehouse due to current purchasing decisions were small since the growth in the inventory levels at the stores was, on average, about 7%. These larger inventories, in both warehouse and stores resulted in an average cost increment of 8%. Nonetheless, the year total cost reduction, in comparison to company current practice, is still about 28%. The small impact in the total costs can be explained by the fact that the larger increase in inventory levels happens at the warehouse, where storage costs are small.

In future work, we aim to investigate with detail the moments of shortage in order to prevent these situations. Possible solutions are: (i) to solve the problem considering a safety stock or (ii) to implement special policies for campaign moments.

References

Allen, S.C. (1958). “Redistribution of total stock over several user locations”. *Naval Research Logistics Quarterly*, 5(4), 337-345.

Bassok, Y., Ernst, R. (1995). “Dynamic Allocations for Multi-Product Distribution”. *Transportation Science*, 29(3), 256–266.

Benito, E. F., & Whybark, D. C. (1986). "Multiple Criteria ABC Analysis", *International Journal of Operations & Production Management*, 6(3), 38-46.

Bertazzi, L., Paletta, G., & Speranza, M. (2005). “Minimizing the Total Cost in an Integrated Vendor-Managed Inventory System”. *Journal of Heuristics*, 11(5–6), 393– 419.

Bradley, S. P., Arnoldo C. H., & Thomas L. M. (1977). “*Applied Mathematical Programming”.* Reading, MA: Addison-Wesley Publishing Company

Brummelle, S. L. and McGill, J. I. (1993). “Airline Seat Allocation with Multiple Nested Fare Classes”. *Operations Research*, 41(1), 127-137.

Chen, M.; Wang; Y., & Hung, T. (2014). “Establishing an order allocation decision support system via learning curve model for apparel logistics”. *Journal of Industrial &Production Engineering*, 31(5), 274-285.

Cretú, B., Fontes, D. B. M. M. “Multi-Product Distribution”. *Proceedings of 2100 Projects Association Joint Conferences 6*, in press 2018

Dauod, H., Wang, H., Khader, N., Yoon, S. W., Srihari, K. (2017). “Real-time Dispenser Replenishment Optimization Based on Receding Horizon Control”. *Procedia Manufacturing*, 11, 1782-1789.

Elf, M., Gutwenger, C., Junger, M., & Rinaldi, G. (2001). “Branch-and-Cut Algorithms for Combinatorial Optimization and Their Implementation in ABACUS”. *Lecture Notes in Computer Science Computational Combinatorial Optimization*, 157-222

Eliiyi, D., Yurtkulu, E. Z., Şahin, D. Y. (2011). “Supply chain management in apparel Industry: a transshipment problem with time constraints”. *Tekstil ve Konfeksiyon*; 21(2), 176-181.

Fortenbacher, P., Ulbig, A., & Andersson, G. (2018). "Optimal Placement and Sizing of Distributed Battery Storage in Low Voltage Grids Using Receding Horizon Control Strategies". *IEEE Transactions on Power Systems*, 33(3), 2383-2394

Gallego, G., & Simchi-Levi, D. (1990). “On the Effectiveness of Direct Shipping Strategy for the One-Warehouse Multi-Retailer R-Systems”. *Management Science*, 36(2), 240-243.

Goodwin, G. C., Seron, M. M., Middleton, R. H., Zhang, M., Hennessy, B. F., & Stone, P. M. (2006). “Receding horizon control applied to optimal mine planning”. *Automatica*, 42(8), 1337–1342.

IBM ILOG. (2009). *IBM ILOG CPLEX V12.1: User's manual for CPLEX*. Available at https://www.ibm.com/support/knowledgecenter/SSSA5P_12.6.2/ilog.odms.studio.help/pdf/usrcplex.pdf. Last accessed on 30.8.2018.

Kang, J. H., Kim, Y. D. (2010). “Coordination of inventory and transportation managements in a two-level supply chain”. *International Journal of Production Economics*, 123 (1) 137-145.

Kumar A., Schwarz L. B., Ward J. E. (1995). “Risk-Pooling Along a Fixed Delivery Route Using a Dynamic Inventory-Allocation Policy”. *Management Science*, 41(2):344-362.

Lambert, A., & Luss, H. (1982). “Production planning with time-dependent capacity bounds”. *European Journal of Operational Research*, 9 (3), 275–280.

Lee, D. J., & Jeong, I. (2010). “A distributed coordination for a single warehouse-multiple retailer problem under private information”. *International Journal of Production Economics*, 125 (1), 190-199.

Luo, K., Bollapragada, R., & Kerbache, L. (2017). “Inventory allocation models for a two-stage, two-product, capacitated supplier and retailer problem with random demand”. *International Journal of Production Economics*, 187, 168-181.

R Core Team. (2014). “R: A language and environment for statistical computing”. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.r-project.org/. Last accessed on 10.5.2018.

Silver, E. A., Meal, H. C. (1973). “A heuristic for selecting lot size quantities for the case of a deterministic time-varying demand rate and discrete opportunities for replenishment”. *Production and Inventory Management: journal of the American Production and Inventory Control Society*, 14(2), 64-74.

Shah, A. K., Oppenheimer, D. M. (2008). “Heuristics made easy: An effort-reduction framework”. *Psychological Bulletin*, 134(2), 207-222.

Speranza, M. G., & Ukovich, W. (1994). “Minimizing Transportation and Inventory Costs for Several Products on a Single Link”. *Operations Research*, 42(5), 879-894.

Thomassey, S. (2010). “Sales forecasts in clothing industry: The key success factor of the supply chain management”. *International Journal of Production Economics*, 128(2), 470-483.

Topaloglu, H. (2005). “An approximate dynamic programming approach for a product distribution problem”. *IIE Transactions*, 37(8), 697–710.

Wagner, H. M., & Whitin, T. M. (1958). “Dynamic Version of the Economic Lot Size Model”. *Management Science*, 5(1), 89-96.

Xu, X., Cai, X., Liu, C., & To, C. (2007) "Optimal Commodity Distribution for a Vehicle with Fixed Capacity Under Vendor Managed Inventory". *Combinatorics, Algorithms, Probabilistic and Experimental Methodologies Lecture Notes in Computer Science*, 329-339.