UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

INTEGRATION OF INVENTORY CONTROL DECISIONS WITH FACILITY LOCATION FOR SEVERAL DEMAND CLASSES

TESIS PARA OPTAR AL GRADO DE DOCTOR EN
SISTEMAS DE INGENIERÍA

PABLO FELIPE ESCALONA RODRÍGUEZ

PROFESOR GUÍA:
FERNANDO ORDOÑEZ PIZARRO

MIEMBROS DE LA COMISIÓN
VLADIMIR MARIANOV KLUGE
PABLO MIRANDA GONZÁLEZ
DENIS SAURE VALENZUELA

SANTIAGO DE CHILE
DICIEMBRE 2015

2

RESUMEN EJECUTIVO

RESUMEN DE LA TESIS PARA OPTAR AL TÍTULO DE:

Doctor en Sistemas de Ingeniería.

POR: Pablo Felipe Escalona Rodríguez

FECHA: 30/12/2015

PROFESOR GUÍA: Fernando Ordóñez Pizarro

INTEGRATION OF INVENTORY CONTROL DECISIONS WITH FACILITY LOCATION FOR SEVERAL DEMAND CLASSES

Esta tesis es acerca de modelos de localización e inventarios con niveles de servicio diferenciados para productos de alta rotación. La motivación detrás de esta tesis es el trabajo desarrollado por mas de 10 años en el Centro Integrado de Manufactura y Automatización de la Universidad Técnica Federico Santa María en Valparaíso, Chile. Durante esos 10 años desarrolle varios trabajos de ingeniería en empresas manufactureras de la Quinta Región (Valparaíso). Todas estas empresas venden sus productos a través de grandes cadenas de retail y también a través de pequeños retails. En estos años he visto como las grandes cadenas de retail en Chile se han concentrado y forzado a sus proveedores a segmentar a sus clientes en función de sus requerimientos de nivel de servicio. Mas paradójico es que estos grandes cadenas de retail exijan altos niveles de servicio y no estén dispuestos a pagar por ello.

En esta tesis estudiamos varias políticas de inventario para proveer niveles de servicio diferenciados en una red de distribución de ítem de alta rotación, por ejemplo, estudiamos las políticas *Separate Stock*, *Round-up* y *critical level*. En el caso de la política *critical level*, no existen trabajos previos que implementen esta política para productos de alta rotación. Por lo tanto, si esta política se expande a una red distribución, sera necesario en una primera etapa desarrollar una formulación teórica para obtener sus parámetros óptimos.

Un breve resumen de las contribuciones de esta tesis se indican a continuación:

- Modelamos y resolvimos dos modelos de nivel crítico. En el primero se considera que el nivel de servicio se mide por la probabilidad de satisfacer toda la demanda de cada clase

durante un ciclo de reabastecimento desde el inventario disponible (nivel de servicio tipo 1), y en el segundo se considera que los backorders de cada clase son penalizados con costos diferenciados;

- Modelamos y resolvimos un problema de diseño de redes de distribución que considera la capacidad de la red de proveer y satisfacer diferentes niveles de servicio usando las siguientes políticas: *single class allocation*, *global round-up*, *separate stock*, *local round-up* and *critical level*.

ABSTRACT

RESUMEN DE LA TESIS PARA OPTAR AL TÍTULO DE:

Doctor en Sistemas de Ingeniería.

POR: Pablo Felipe Escalona Rodríguez

FECHA: 30/12/2015

PROFESOR GUÍA: Fernando Ordóñez Pizarro

INTEGRATION OF INVENTORY CONTROL DECISIONS WITH FACILITY LOCATION FOR SEVERAL DEMAND CLASSES

This thesis is about location-inventory models with differentiated service levels for quick turnover items. The motivation behind this thesis is the work that I developed over 10 years in the Integrated Center of Manufacturing and Automation of the Universidad Técnica Federico Santa María in Valparaíso, Chile. During those 10 years I developed several engineering projects in manufacturing companies in the region of Valparaiso. All these companies sell items with high demand volume through large retail chains and through small retailers. In these years I have seen large retail chains in Chile have been concentrated and forced these companies to segment their customers based on the level of service. Most paradoxical is that these large retail chains require high service levels for which they are not willing to pay.

In this thesis we studied several types of inventory control policies to provide differentiated service levels in a distribution network of fast-moving items, e.g., *Separate Stock*, *Round-up* and *critical level* policies. In the case of critical level policy, to the best of our knowledge, there does not exist previous work implementing this policy when demand volume is large. Therefore, if this policy is to be extended to a distribution network, it is first necessary to develop a theoretical formulation to obtain its optimal parameters.

Let us shortly sum up the contributions in this thesis.

- We model and solve two critical level model for quick turnover items. The first one consider that the service level is measured by the probability of satisfying the entire demand of each class during a replenishment cycle from the on-hand inventory (service level type 1), and the second one considers that the backorders are penalized with differentiated costs;

- We model and solve a supply chain design problem that considers the ability of the distribution network, to provide and fulfill different service levels using *single class allocation*, *global round-up*, *separate stock*, *local round-up* and *critical level* policies.

For

Claudia, Julieta and Lautaro

Acknowledgment

Several people have contributed to this thesis. First of all, I thank my supervisor Fernando Ordóñez for outstanding support and encouragement through the 3 years. I also thank Professor Vladimir Marianov for his support. Most of the ideas presented in the thesis were conceived during long discussions with Fernando and Vladimir. For this I am deeply indebted.

I want to thank Edgard Iturrieta and Ismael Kauak for their help and contributions to Chapter 4 and Carlos Reyes for their contribution in Chapter 6.

Finally, I thank my family, Claudia, Julieta, and Lautaro, and, in particular, my wife for her support.

Contents

List of Tables

List of Figures

1 | Introduction

The fulfillment of required service levels for different classes of customers in terms of product availability is an important issue when designing a distribution network. Especially when service level failure can lead to loss of customers or fines. The optimal design of the distribution network should not only determine the number of distribution centers (DC) to locate (*stocking locations*), where to locate, what kinds of customers should be assigned to each DC and how much inventory to keep each of them, but it also should prescribe how to meet the required service level for each class or category of customer demand.

Models that integrate simultaneously inventory and location decisions (Daskin et al. (2002), Shen et al. (2003), Shen (2005), Shen and Daskin (2005), Miranda and Garrido (2004), Snyder et al. (2007), Ozsen et al. (2008), You and Grossmann (2008), Atamtürk et al. (2012)), consider that the distribution network is dominated by continuous review (Q, r) policy, full backorder, deterministic lead time and normally distributed demand in which the same service level is provided to the whole network, i.e., they do not integrate the requirement of different demand classes in the optimal configuration of the distribution network. Considering that the entire network is balanced in terms of service level is not always realistic.

Several types of inventory policies can be implemented to provide differentiated service level to the distribution network. In this thesis we propose classifying the policies that provide differentiated service levels into two types. The first group of policies imposes general service conditions over the entire network distribution. The simplest mechanism is that each DC serves a single demand class (*single class allocation*). This policy tends to increase the number of DCs in the network and not to take advantage of the risk pooling benefits. Another mechanism is to set the service level of the entire distribution network based on a preset level corresponding to the highest priority class (*global round-up policy*). The second type of policy imposes conditions on the operation of the inventory system at each DC. In this case, the simplest mechanism is to impose that each

DC serves the demand assigned to it from a common stockpile and uses separate safety stocks for each class (*separate stock policy*). However, separating the safety stocks in each DC does not take advantage of the benefit of centralized inventories. The separate stock policy can be outperformed imposing at each DC, a mechanism that serves all demand assigned to it from a common stockpile and sets the safety stock as the maximum required between the sets of classes assigned to it (*local round-up policy*). In this case, although all demand of a DC is centralized and the variability reduced, this policy may provide too much inventory for classes that require less service level than the maximum. A third mechanism of this type consists in each DC serving the demand assigned to it from a common stockpile, but using a *critical level policy* for rationing the inventory between different classes. With this policy, as soon as the inventory level falls below a critical level, the low priority demands are not attended. Its main application is in inventory systems that must provide differentiated service levels to two or more classes of demand. This policy can be implemented for several ordering and review policies. For example, a traditional (Q, r) model is extended using a critical level policy to a $(Q, r, \mathbf{C})$ inventory model, where Q is the fixed batch size, r is the reorder point and $\mathbf{C} := \{C_1, ..., C_{n-1}\}$ denote a set of critical levels for rationing n classes of demand (Nahmias and Demmy (1981); Melchiors et al. (2000); Deshpande et al. (2003); Isotupa (2006); Arslan et al. (2007); Wang et al. (2013a)), and $(S-1, S)$ policies are extended to a $(S-1, S, \mathbf{C})$, where S denotes the base stock level (Ha (1997a,b, 2000); De Vericourt et al. (2000, 2002); Bulut and Fadiloğlu (2011); Piplani and Liu (2014) for make-to-stock production system, and Dekker et al. (2002, 1998); Möllering and Thonemann (2010); Fadıloğlu and Bulut (2010); Wang et al. (2013b) for lot-for-lot inventory systems).

Except for the critical level policy, all the mechanisms for providing differentiated service levels operating under continuous review policy and normally distributed demand may extend to distribution networks. In general, the complexity is not in its formulation, if not the method of resolution due to the non-linearity inducing inventories in the network configuration.

Assuming normally distributed demand, previous inventory-location models, explicitly or implicitly assume that the distribution network deals with quick turnover items, i.e., products with high demand volume. Examples include non-perishable food, toiletries, over-the-counter drugs, cleaning supplies, building supplies and office supplies. The distribution channels of these products have been concentrated in large retails chains requiring high service level in terms of product availability at the supplier‘s expense. Therefore, many wholesalers segment their customers based on service level. The simplest segmentation is to classify customers into two demand classes, (i) high priority class will correspond to large retail chains that require high service levels and (ii) low priority class corresponding to small retailers which can be provided lower service levels.

Let us now consider the implementation of a critical level policy for items with high demand volume (fast-moving items or items with high rotation rate). For these items, it is usually more convenient and efficient to model the demand over a time period by a continuous distribution, e.g., normal or gamma distributions (Axsäter (2007); Ramaekers and Janssens (2008)). Previous work on critical level policy has only considered the case of discrete demand, in particular Poisson distributed demand, which is how the demand for slow-moving items is modeled. Therefore, to the best of our knowledge, there does not exist previous work that implements a critical level policy when demand volume is large.

In summary, to integrate the requirements of different classes of customers in terms of product availability on the optimal configuration of a distribution network for fast moving items, it is necessary:

- develop a theoretical formulation of the critical level policy with continuous review and continuous distribution demand, in particular normal distribution, and then integrate it into a location-inventory model,
- model and solve three simple ways to provide differentiated service levels in a distribution network, i.e., *local round up*, *separate stock* and *single class allocation* policies.

The objective of this thesis is to integrate the requirements of different classes of customers, in terms of product availability, on the optimal configuration of a distribution network with fast-moving items. Furthermore, we have the following specific objectives:

- determine the optimal parameters of a continuous review (Q, r, C) policy for fast-moving items when rationing is used to provide differentiated service levels to two demand classes (high and low priority) and the service level is measured by the probability of satisfying the entire demand during a replenishment cycle from on-hand inventory (service level type 1),
- determine the optimal parameters of a continuous review (Q, r, C) policy for fast-moving items when the inventory system facing random demand of two customer classes (high and low priority) and backorders of each class have different penalty costs,

- modeling and solving a supply chain design problem for fast-moving items that considers the ability of the distribution network to provide and fulfill different service levels using critical level, separate stock, single class allocation, global round-up and local round-up policies.

The remainder of this thesis is structured as follows:

A review of related work in location-inventory models and the critical level policy is discussed in the chapter 2.

In chapter 3 we analyze the use of a constant critical level policy for fast-moving items where rationing is used to provide differentiated service levels. In this chapter we consider a continuous review (Q, r, C) policy with two demand classes, the service level for each demand class is measured by service level type 1 and to determine the operational characteristics of the inventory system we use a *hitting time* approach. We also consider the *threshold clearing mechanism* of Deshpande et al. (2003) to allocate backorders when multiple outstanding orders exist. The problem assumes a fast-moving item which makes it reasonable to model the demand over a time period by a continuous distribution with positive support. Given the inventory control strategy, we formulate a non-linear problem with chance constraints, denoted (**SLP**), to determine the parameters of the critical level policy. We propose to solve a relaxation of (**SLP**) which is able to provide good bounds. For strictly increasing non-negative demand, we characterize the optimal solution of this relaxation through a system of equations. We further extend this solution, under mild assumptions, when the normal distribution is used as an approximation of the non-negative demand.

In chapter 4 we analyze the constant critical level policy for fast-moving items when the inventory system facing random demand of two customer classes (high and low priority). The inventory system operates under continuous review (Q, r) policy, with a critical threshold value C, full-backorder and deterministic lead time. Penalty cost of backorders of high priority class are greater than the low priority class and demand of each class is characterize by a strictly increasing non-negative demand. We also characterize the demand of each class with a normal distribution, as a approximation of the non-negative demand. Given the inventory control strategy and using the *state-dependent demand* approach to model the operation of the inventory system, in which a threshold mechanism is adopted to allocate backorders when multiple outstanding orders exist, we propose an approximate expression for the on-hand inventory based on a convex approximation of backorders for each class. The approximation consider that the demand class during the rationing period is proportional to the total demand for both classes during this period. This approach allows us to formulate a convex cost minimization problem to determine the parameters of the critical level policy, that can be solved through a equations system derived from the KKT conditions or using a convex non linear solver.

In chapter 5 we analyze the design of a distribution network able to provide differentiated service levels in terms of product availability for two demand classes (high and low priority). To provide differentiated service levels we consider a critical level policy, and that the service level provided by a DC is measured by the probability of satisfying the entire demand of each class assigned to the DC during a replenishment cycle from on-hand inventory. We formulate the location-inventory model with differentiated service levels, denoted (**P0**), as an MINLP problem with chance constraints and nonlinear objective function. The chance constraints of (**P0**) correspond to the service levels constraints. We observe that the location-inventory model with a single service level is a relaxation of (**P0**). We reformulate the location-inventory model with a single service level as a conic quadratic mixed integer program from which we obtain a lower bound of (**P0**). Using the resulting configuration of the relaxation of (**P0**) in terms of location and allocation variables we obtain the optimal control parameters of the critical level policy at each DC. The result is an upper bound (feasible solution) for the problem (**P0**). Furthermore, we propose a method to improve the solution based in the risk pooling effect.

In chapter 6 we study the *single class allocation*, *separate stock* and *local round-up* policies to design a distribution network able to provide differentiated service levels for two demand classes (high and low priority). For each policy, we formulate an integer non-linear problem (INLP). We show how to formulate the *single class allocation*, *separate stock* and *local round-up* problems as conic quadratic mixed-integer problems. In particular, *local round-up* policy present particular challenges in their formulation, for which we propose a Lagrangian relaxation over the conic quadratic mixed-integer formulation. We compare the performances of using the five inventory policies for different parameters settings.

Chapter 7 correspond a illustrative example of the location-inventory models with differentiated service levels developed in chapters 5 and 6. The industrial application correspond a company that manufactures products derived from fruits which requires determining the number of distribution centers (DC) to locate in Santiago (Chile), where to locate, what kinds of customers should be assigned to each DC, how much inventory to keep each of them, and how to meet the required service level of their customer. Our results indicates that the lower cost configuration is achieved with the critical level policy.

Finally, Chapter 8 contains the final conclusions of this thesis.

2 | Related work

In this chapter, a review of related works of location-inventory models and the critical level policy is discussed.

2.1 Location-inventory models

The traditional structure of Facility Location Problem (Erlenkotter (1978)) does not consider the relationship between location and inventory control decisions, nor its impact on the distribution network configuration. This is because the distribution network design is solved sequentially, by first solving the location problem and then the inventory problem. This is related to the natural separation between strategic and tactical decision making. However, when these decisions are addressed separately, it often results in suboptimal solutions. In the last decade there has been a strong move towards integrated models of inventories and location. These models simultaneously determine the location of the DCs that will be opened, the allocation of customers to DCs and the optimal parameters of the inventory policy so as to minimize the total system cost. A comprehensive characterization in location-inventory models can be found in Sadjadi et al. (2015).

Our work focuses on location inventory models that integrate the service level, in terms of product availability, in its formulation. In this sense, Daskin et al. (2002) study a location-inventory model that incorporates fixed facility location cost, ordering, holding and safety-stock inventory cost at the DCs, transportation costs from the supplier to the DCs, and local delivery costs from the DCs to the customers. The main difficulty of this model is that the inventory costs at each DC are not linear respect to customer assignments. The model is formulated as a nonlinear integer program and solved by Lagrangian relaxation for a special case in which the ratio between the variance and expected demand is constant for all customers. Shen et al. (2003) analyze the same problem as Daskin et al. (2002). Their work restructures the model into a set-covering integer programming model and use column generation to solve the LP-relaxation of the set covering model.

The model of Daskin et al. (2002) and Shen et al. (2003) has been generalized in different directions. For example: Shen (2005) generalizes the model to a multi-commodity case with a general cost function and proposes a Lagrangian-relaxation solution algorithm. Shen (2005) also relaxes the assumption that the variance of the demand is proportional to the mean for all customers and proposes a Lagrangian-relaxation approach using an algorithm proposed by Shu et al. (2005). Shen and Daskin (2005) introduce a service level element in the model through the distance coverage and propose a weighting method and a heuristic solution approach based on genetic algorithms. Snyder et al. (2007) present a stochastic version of the model. Ozsen et al. (2008) study a capacitated version of the model. Miranda and Garrido (2004) also study a capacitated version of the model and propose a Lagrangian-relaxation solution algorithm. You and Grossmann (2008) relax the assumption that each customer has identical variance-to-mean ratio, reformulating the INLP model as a MINLP problem and solve it with different solution approaches, including a heuristic method and a Lagrangean relaxation algorithm. Atamtürk et al. (2012) also relax the assumption that each customer has identical variance to mean ratio and reformulate the INLP model of Daskin et al. (2002) as conic quadratic mixed-integer problem and added cuts to improve the computational results. They consider cases with uncapacitated facilities, capacitated facilities, correlated retailer demand, stochastic lead times, and multi commodities. Atamtürk et al. (2012) show, through a computational study, that the conic formulation outperforms the column generation and Lagrangian based methods considered up to now. Shahabi et al. (2014) study a capacitated version with correlated retailer demand and propose a solution approach based on an outer approximation strategy.

All of the above authors assume that the inventory system at each DC operates under a continuous review (Q,r) policy with type I service level and full-backorder. Under this policy, a replenishment order Q is emitted when the inventory level falls below the reorder point r. Based on the results of Axsäter (1996) and Zheng (1992), previous work has approximated the (Q,r) model assuming that each DC determines the replenishment batch Q using an EOQ model and determines the reorder point r and the safety stock ensuring that the probability of a stockout at each DC is less than or equal to some preset service level. This preset service level is the same for all the distribution network. Further, a normally distributed demand is assumed at each DC as an approximation for a high volume Poisson demand process. With these approximations, the parameters for the continuous review (Q,r) policy are the result of the optimal allocation of customers to DCs.

Our work focuses on situations where customers may require different service levels or that there are different demand classes, which can be more realistic in many cases. To the best of our knowledge, there does not exist previous works that integrates differentiated service levels in the optimal configuration of the network distribution.

Several types of inventory control policies can be implemented in a distribution network to deal with different service requirements. We propose classifying this policies into two types. The first group of policies imposes general service conditions over the entire network distribution. The simplest mechanism is that each DC serves a single demand class, to which we refer as *single class allocation*. This policy tends to increase the number of DCs in the network and not to take advantage of the risk pooling benefits (Eppen (1979)). Another mechanism is to set the service level of the entire distribution network based on a preset level corresponding to the highest priority class, to which we refer as *global round-up policy*. This policy tends to provide too much inventory for classes that require less service level than the maximum. The second type of policy imposes conditions on the operation of the inventory system at each DC. In this case, the simplest mechanism is to impose that each DC serves the demand assigned to it from a common stockpile and uses separate safety stocks for each class (*separate stock policy*). However, separating the safety stocks in each DC does not take advantage of the benefit of centralized inventories. The separate stock policy can be outperformed imposing at each DC, a mechanism that serves all demand assigned to it from a common stockpile and sets the safety stock as the maximum required between the sets of classes assigned to it (*local round-up policy*). In this case, although all demand of a DC is centralized and the variability reduced, this policy may provide too much inventory for classes that require less service level than the maximum. A third mechanism of this type consists in each DC serving the demand assigned to it from a common stockpile, but using a *critical level policy* for rationing the inventory between different classes. With this policy, as soon as the inventory level falls below a critical level, the low priority demands are not attended.

In the current thesis we focus on critical level policy because as well as using the advantage of the pooling effect, it has the flexibility of providing different service levels to different customer classes without provide too much inventory for classes that require less service level than the maximum or increase the number of DCs in the network.

2.2 Critical level policy

A comprehensive review of inventory rationing can be found at Kleijn and Dekker (1999) and a classification at Teunter and Haneveld (2008). In particular, Kleijn and Dekker (1999) classified inventory systems subject to multiple classes of demand based on the *review policy* (continuous and periodic) and the *number of classes* (2 or n classes). The above classification is extended by Teunter and Haneveld (2008), incorporating *shortage treatment* (backorder or lost sale), *rationing policy* (no-rationing, static, dynamic), the *ordering policy* and the way that the time is modeled (discrete or continuous).

Our model corresponds to a constant critical level (Q,r,C) policy of continuous review and demand. In this sense, Nahmias and Demmy (1981) were the first ones that studied the continuous review policy with two demand classes. They assumed a (Q,r,C) policy, Poisson demand, full-backorders and deterministic lead time. This work does not determine the optimal parameters of the critical level policy, but develops an approximate expression for the expected backorder per cycle for both demand classes when there is at most one outstanding order and uses the *hitting time* to model the inventory behavior. Melchiors et al. (2000) also analyze a (Q,r,C) inventory model, deterministic lead time and two demands classes, but unlike Nahmias and Demmy (1981), they consider a lost sales environment. In order to determine the optimal parameters of the critical level policy these authors propose a cost optimization problem and present a numeric procedure for its resolution. They assumed Poisson demand and used the *hitting time* and renewal theory to operationally characterize the inventory system. Isotupa (2006) presents a model with the same assumptions as Melchiors et al. (2000) but with exponentially distributed lead time.

When implementing a continuous review (Q,r,C) critical level policy with full backorder, it may happen that the incoming replenishment batch is not large enough to cover the backorders. Therefore, it is important how the backorders of the different classes are satisfied. According to Möllering and Thonemann (2010) it is optimal to fill backorders from high priority classes first when dealing with penalty costs. This form of clearing the backorders is called *priority clearing mechanism*. This policy is difficult to analyze mathematically and given its complexity the literature has focused on manageable but sub-optimal rules, e.g., the *threshold clearing mechanism* from Deshpande et al. (2003) and the *FCFS type clearing scheme* from Arslan et al. (2007). Deshpande et al. (2003) analyzed the same rationing model as Nahmias and Demmy (1981), but without restricting the number of outstanding orders. They derived expressions for the average backorders per cycle and for the expected steady-state on-hand inventory and backorder using a state-dependent demand approach. Based on these expressions, Deshpande et al. (2003) proposed a cost optimization model and developed algorithms to compute the optimal parameters of the critical level policy. Arslan et al. (2007) presents a service level model to obtain the optimal parameters of a critical level policy with multiple demand classes under the assumptions of Poisson demand, deterministic lead time, and a continuous-review (Q,r) policy. Wang et al. (2013a) analyzed the rationing policy under the same operational conditions than Deshpande et al. (2003), but considered a mixed service criteria with penalty costs and service level constraints (fill-rate). In that work, they show numerically that the priority clearing mechanism does not always outperform the threshold clearing mechanism when dealing with service levels constraints.

In certain situations a dynamic rationing policy, which allows the critical level to change based on the number and ages of outstanding orders, can outperform a constant critical level policy (Q,r,C). Fadiloglu and Bulut (2010) examine a dynamic rationing policy, in a continuous review (Q,r) inventory model with Poisson demand and deterministic lead time. The authors use simulation-based approaches to find efficient solutions for the cases with backordering and lost sales.

From the literature review conducted only Dekker et al. (2002), Arslan et al. (2007), Wang et al. (2013b) and Möllering and Thonemann (2010) use a service level problem approach to determine the optimal parameters of the critical level policy. These four articles consider the same service level problem: to minimize the expected on-hand inventory subject to having the service level provided to each class exceed its preset level. Depending on the operating conditions defined for the inventory system, what varies is the formulation of the inventory on-hand value and the service level provided to each class. Dekker et al. (2002) analyzed the critical level policy when the inventory system works under a continuous review lot-for-lot policy, lost sales and Poisson demand. These authors derive expressions for fill-rate and present an efficient method to obtain optimal solutions. Möllering and Thonemann (2010) analyze a periodic review base-stock policy with two demand classes, deterministic lead time, discrete demand distribution and full backorder. That work models the inventory system as a multidimensional Markov chain and optimally solves a service level problem, based on a service level of type 1 and another on fill-rate. Wang et al. (2013b) analyzed the same model as Möllering and Thonemann (2010), but considered an anticipated rationing policy.

This policy reserves inventory for the high priority classes considering a constant critical level and the coming replenishment of the next period.

In summary, previous research on inventory rationing solved periodic or continuous review problems with discrete demand. Therefore, to the best of our knowledge, there is no constant critical level model for the case of continuous demand distribution considered in this thesis.

3 | Critical Level Rationing In Inventory Systems With Continuously Distributed Demand

This chapter analyzes the use of a constant critical level policy for fast-moving items where rationing is used to provide differentiated service levels to two demand classes (high and low priority). Previous work on critical level models, with either a continuous or periodic review policy, has only considered slow-moving items with Poisson demand. In this chapter we consider a continuous review (Q, r, C) policy with two demand classes that are modeled through continuous distributions and the service levels are measured by the probability of satisfying the entire demand of each class during the lead time. We formulate a service level problem as an non-linear problem with chance constraints for which we optimally solve a relaxation obtaining a closed form solution that can be computed easily. For instances we tested, computational results show that our solution approach provide good-quality solutions that are on average 0.3% from the optimal solution.

3.1 Service level problem for strictly increasing non-negative demand

Consider a facility that holds inventory of a single type of product to serve two demand classes $i = 1, 2$, where class 1 is high priority and class 2 is low priority. Let $D_i(t, t+\tau)$ be the total demand of class i in the interval $(t, t+\tau]$, and $D(t, t+\tau) = D_1(t, t+\tau) + D_2(t, t+\tau)$ the total demand of both classes in the interval $(t, t+\tau]$. We denote by $F_{D_i(\tau)}(x)$ the cumulative distribution function of the total demand of class i in $[0, \tau]$ and $F_{D(\tau)}(x)$ the cumulative distribution function of the total demand of both classes in $[0, \tau]$.

In this chapter we consider fast-moving items for which is more representative and efficient to model the demand over a time period by a continuous distribution. Following Zheng (1992) we assume that the total demand of each class are represented by a non-decreasing stochastic process with continuous sample paths, and stationary and independent increments. For simplicity, we will refer to this as strictly increasing non-negative demand. This is a common assumption in stochastic inventory models (Axsäter (2007)) and is implicitly assumed in most (elementary) textbooks on inventory management. However, the assumptions of independence and continuity are conflicting; therefore, rigorously speaking, the assumption is approximate (Browne and Zipkin (1991)). Note that under stationary and independent increments, $D_i(\tau) := D_i(0, \tau) = D_i(t, t + \tau)$ for any $t \geq 0$, $i = 1, 2$.

Inventory is replenished according to a continuous review (Q, r, C) policy that operates as follows. When the inventory position falls below a reorder level r, a replenishment order for Q units is placed and arrives a fixed $L > 0$ time units later. Demand from both classes are filled as long as the on-hand inventory level is greater than the critical level C, otherwise only high priority demand is satisfied from inventory on-hand and low priority demand is backordered. If on-hand inventory level reaches zero both demands are backordered. To clear backlogged orders, we consider the threshold clearing mechanism of Deshpande et al. (2003).

Given the inventory control strategy, our objective is to find the parameters of the critical level policy that minimize the sum of ordering and holding costs per unit time subject to satisfying the required service level for each class. In this chapter, the service level is measured by the probability of satisfying the entire demand of each class during the lead time from on-hand inventory (service level type 1), which does not depend of the replenishment batch quantity. Let $\alpha_i(r, C)$ be the provided service level to class i and $\overline{\alpha}_i$ the preset service level for class i, where $\overline{\alpha}_1 > \overline{\alpha}_2 > 0$. Then, the service level problem is:

$$\min_{Q,r,C} \quad AC(Q, r, C) \tag{3.1}$$

$$\text{s.t:} \quad \alpha_i(r, C) \geq \overline{\alpha}_i \qquad \forall i = 1, 2 \tag{3.2}$$

$$Q, r, C \geq 0. \tag{3.3}$$

where $AC(Q, r, C)$ is the average cost per unit time, i.e., the sum of ordering and holding costs per unit time. To develop expressions for $\alpha_i(r, C)$, $i = 1, 2$, and $AC(Q, r, C)$ we use a hitting time approach as in Nahmias and Demmy (1981) and the threshold clearing mechanism of Deshpande et al. (2003) to allocate backorders when multiple outstanding orders exist.

The hitting time $\tau^x_{H,D}$ is defined as the amount of time that elapses until the demand D reaches x for the first time, i.e.,

$$\tau^x_{H,D} = \inf\{\tau > 0 \mid D(\tau) > x\}. \tag{3.4}$$

Since we assume strictly increasing non-negative demand, we have $\mathbb{P}(\tau^x_{H,D} \leq \tau) = \mathbb{P}(D(\tau) \geq x)$. Therefore, the distribution function of the hitting time $\tau^x_{H,D}$, for a fixed $x > 0$ is $F^x_{H,D}(\tau) = 1 - F_{D(\tau)}(x)$, and its density distribution is:

$$f^x_{H,D}(\tau) = -\frac{\partial F_{D(\tau)}(x)}{\partial \tau}. \tag{3.5}$$

Many authors have discussed the hitting time process for strictly increasing non-negative demand. However, an explicit expression for the density of hitting time is not possible in many cases. Meerschaert and Scheffler (2008) develop a density formula for the hitting time of any strictly increasing non-negative demand based on the Laplace transform of the hitting time. Park and Padgett (2005) derived a exact density distribution of hitting time for a gamma process using the same procedure described by equation (3.5).

3.1.1 Average cost per unit time.

Let μ be the total average demand per unit of time, h be the holding cost per unit and unit time and S the ordering cost. Then the average cost per unit time is $AC(Q, r, C) = S\frac{\mu}{Q} + h\mathbb{E}(OH(\infty))$, where $OH(\infty)$ is the steady-state on-hand inventory (Axsäter (2007)).

In a (Q, r, C) policy with full-backorders and deterministic lead time, the inventory level is the on-hand inventory net of all backorders, i.e., $IL(t + L) = OH(t + L) - B_1(t + L) - B_2(t + L)$, where $IL(t + L)$ denotes the inventory level, $OH(t + L)$ denotes on-hand inventory and $B_i(t + L)$ denotes class i backorders, $i = 1, 2$, all at time $t + L$. Furthermore, for a (Q, r, C) policy with full-backorders and deterministic lead time it is still valid that $IL(t + L) = IP(t) - D(L)$, where $IP(t)$ denotes the inventory position at time t (Deshpande et al. (2003)). Under strictly increasing non-negative demand, $IP(t)$ will be uniformly distributed on $(r, r + Q]$ in steady state and independent of lead time demand (Zheng (1992) refers to Serfozo and Stidham (1978) and Browne and Zipkin

(1991) for a detailed discussion of this assumption). Then, the on-hand inventory at time $t + L$ is $OH(t+L) = IP(t)-D(L)+B_1(t+L)+B_2(t+L)$. Taking expected value and limit $t \to \infty$, the expected on-hand inventory at steady-state is $\mathbb{E}(OH(\infty)) = \frac{Q}{2}+r-\mu L+\mathbb{E}(B_1^\infty(Q,r,C))+\mathbb{E}(B_2^\infty(Q,r,C))$, where $\mathbb{E}(B_i^\infty(Q,r,C))$ is the class i steady-state backorder, $i = 1, 2$. Then, the average cost per unit time is:

$$AC(Q,r,C) = S\frac{\mu}{Q} + h\Big(\frac{Q}{2} + r - \mu L + \mathbb{E}(B_1^\infty(Q,r,C)) + \mathbb{E}(B_2^\infty(Q,r,C))\Big) \tag{3.6}$$

We now develop expressions for the backorders of the low and high priority class in steady state using a hitting time approach, the inventory position and the threshold clearing mechanism. We first describe how the inventory system behaves under rationing, and the threshold clearing mechanism of Deshpande et al. (2003).

Consider an arbitrary time $t + L$. By definition, there is rationing at time $t + L$ when $C > OH(t + L) \geq IL(t + L) = IP(t) - D(t, t + L)$. Using the hitting time $\tau_{H,D}^{IP(t)-C}$ defined in equation (3.4), this last condition states that if there is rationing at $t + L$ then $\tau_{H,D}^{IP(t)-C} < L$. Note that $\tau_{H,D}^{IP(t)-C}$ corresponds to the time required for $IP(t) - C$ demands.

Define t_c as the first time after t when $IP(t) - C$ demand is observed, that is, $t_c = t + \tau_{H,D}^{IP(t)-C}$. If rationing ocurrs at $t + L$ then we have that $\tau_{H,D}^{IP(t)-C} < L$. The threshold clearing mechanism of Deshpande et al. (2003) only comes into play when backorders exist on arrival of a replenishment order and uses t_c to separate which backorders need to be cleared once the replenishment order arrives. The general rules to clear the backorders when the replenishment order arrives are:

1. If the entering replenishment batch is large enough to clear all the backorders and leave the on-hand inventory level above C, then clear all backorders,

2. Otherwise:

   2.1 Clear all backlogged demand that arrived before t_c in the order of arrival (FCFS),

   2.2 Clear any remaining backlogged class 1 demands using FCFS until either all class 1 backorders are filled, or no on-hand inventory remains,

   2.3 Carry over (i.e. continue backlog) all class 2 demands that arrive after t_c.

Note that rule 1 ensures that $OH(t) = IL(t)$ when $OH(t) \geq C$. Rule 2.2 and 2.3 mean that all remaining backorders that cannot be fulfilled by the entering replenishment batch, are carried over to be satisfied in the following replenishment arrivals.

Using the hitting time definition it easy to show that the backorders of both low- and high-priority at time $t + L$, deduced by Deshpande et al. (2003), are respectively:

$$B_2(t+L) = \begin{cases} D_2(L - \tau_{H,D}^{IP(t)-C}) & \text{if } \tau_{H,D}^{IP(t)-C} < L \\ 0 & \sim \end{cases} \tag{3.7}$$

$$B_1(t+L) = \begin{cases} D_1(L - \tau_{H,D}^{IP(t)-C} - \tau_{H,D_1}^{C}) & \text{if } \tau_{H,D}^{IP(t)-C} + \tau_{H,D_1}^{C} < L\,, \\ 0 & \sim \end{cases} \tag{3.8}$$

where $\tau_{H,D_1}^{C} = \inf\{\tau > 0 \mid D_1(\tau) > C\}$ corresponds to the time required for C demands of class 1. The equivalence between equations (3.7), (3.8) and the expressions developed by Deshpande et al. (2003), are given by the fact that: $D_2(t_c, t+L) = D_2(L - \tau_{H,D}^{IP(t)-C})$ and $[D_1(t_c, t+L) - C]^+ = D_1(L - \tau_{H,D}^{IP(t)-C} - \tau_{H,D_1}^{C})$.

Taking expectation of equations (3.7) and (3.8) and conditioning on the inventory position $IP(t)$, the expected backorders at steady state of class 1 and 2 are respectively:

$$\mathbb{E}(B_2^{\infty}(Q,r,C)) = \frac{1}{Q}\int_r^{r+Q}\int_0^L \mathbb{E}(D_2(L-\tau)) f_{H,D}^{y-C}(\tau)\, d\tau\, dy\,, \tag{3.9}$$

$$\mathbb{E}(B_1^{\infty}(Q,r,C)) = \frac{1}{Q}\int_r^{r+Q}\int_0^L \mathbb{E}(D_1(L-\tau)) (f_{H,D}^{y-C} * f_{H,D_1}^{C})(\tau)\, d\tau\, dy\,, \tag{3.10}$$

where: $\mathbb{E}(D_i(L-\tau)) = \int_{-\infty}^{\infty} x f_{D_i(L-\tau)}(x)dx$; $f_{H,D}^{y-C}(\tau) = -\frac{\partial F_{D(\tau)}(y-C)}{\partial \tau}$; $f_{H,D_1}^{C}(\tau) = -\frac{\partial F_{D_1(\tau)}(C)}{\partial \tau}$; and we denote by $f_{H,D}^{y-C} * f_{H,D_1}^{C}(\tau) = \int_0^{\tau} f_{H,D}^{y-C}(\tau - t) f_{H,D_1}^{C}(t)dt$ the convolution of $f_{H,D}^{y-C}(\tau)$ and $f_{H,D_1}^{C}(\tau)$.

3.1.2 Service level type I under rationing policy

We now develop expressions for $\alpha_i(r,C)$ of class $i = 1,2$ using the hitting time approach. We first describe the events to fully meet the demand of each class during the lead time under strictly increasing non-negative demand.

The conditions to fully meet the demand for class 2 in the lead time, under non-negative demand, are that: (i) there does not exist rationing, i.e., $\tau_{H,D}^{r-C} > L$, where $\tau_{H,D}^{r-C}$ is defined in equation (3.4) and corresponds to the time required for $r - C$ demands, or (ii) rationing occurs and there is no class 2 demand, i.e., $\tau_{H,D}^{r-C} < L$ and $D_2(\tau) = 0$, $\forall \tau \in [\tau_{H,D}^{r-C}, L]$. Since $D_2(\tau)$ is defined as strictly increasing non-negative demand, the probability that rationing occurs and there is no demand of the class 2 during this period is zero. Therefore, the service level provided to the low priority class is:

$$\alpha_2(r, C) = \mathbb{P}(D(L) \leq r - C) = F_{D(L)}(r - C). \tag{3.11}$$

The conditions to fully meet the demand of class 1 in the lead time, under non-negative demand, are that: (i) rationing does not exist or (ii) rationing occurs and the class 1 demand during this period not reach the critical level C, i.e., $\tau^{r-C}_{H,D} < L$ and $\tau^{C}_{H,D_1} \geq L - \tau^{r-C}_{H,D}$. Therefore, the service level provided to the high priority class is:

$$\alpha_1(r, C) = \mathbb{P}(D(L) \leq r - C) + \mathbb{P}(D_1(L - \tau^{r-C}_{H,D}) \leq C \ \cap \ \tau^{r-C}_{H,D} < L), \tag{3.12}$$

because, $\mathbb{P}(\tau^{C}_{H,D_1} > L - \tau^{r-C}_{H,D} \ \cap \ \tau^{r-C}_{H,D} < L) = \mathbb{P}(D_1(L - \tau^{r-C}_{H,D}) \leq C \ \cap \ \tau^{r-C}_{H,D} < L)$. Conditioning on the hitting time $\tau^{r-C}_{H,D}$, the service level provided to the high priority class can be expressed as:

$$\begin{aligned} \alpha_1(r, C) &= \int_0^L \mathbb{P}(D_1(L - \tau) \leq C) \ f^{r-C}_{H,D}(\tau) \ d\tau + \mathbb{P}(D(L) \leq r - C) \\ &= \int_0^L \mathbb{P}(D_1(L - \tau) \leq C) \ f^{r-C}_{H,D}(\tau) \ d\tau + \alpha_2(r, C). \end{aligned} \tag{3.13}$$

Note that equation (3.13) verifies that $\alpha_1(r, C) \geq \alpha_2(r, C)$.

Under strictly increasing non-negative demand, the definition of the hitting time $\tau^{r-C}_{H,D}$ implies that reorder point is strictly greater than the critical level, i.e., $r > C \geq 0$. Otherwise the (Q, r, C) policy is not interesting because the provided service level to low priority class is zero. For example, if $r = C \geq 0$ in every lead time exist rationing and the only possibility to fully meet the demand for class 2 is that there is no class 2 demand during the lead time (in this case the lead time is equal to rationing period for class 2). Then, under strictly increasing non-negative demand and $r = C \geq 0$, $\alpha_2(r, r) = \mathbb{P}(D_2(L) \leq 0) = 0$. In the same way, we can conclude that under strictly increasing non-negative demand, for any $C > r \geq 0$, $\alpha_2(r, C) = 0$. Therefore, in this chapter we will study only the case where $r > C \geq 0$.

3.1.3 Problem formulation

Using equation (3.11) and (3.13) we can write the service level problem for strictly increasing non-negative demand, denoted (**SLP**), as the following optimization problem.

Problem (**SLP**):

$$\min_{Q,r,C} \quad S\frac{\mu}{Q} + h\left(\frac{Q}{2} + r - \mu L + \mathbb{E}(B_1^{\infty}(Q,r,C)) + \mathbb{E}(B_2^{\infty}(Q,r,C))\right) \tag{3.14}$$

$$\text{s.t:} \quad \int_0^L \mathbb{P}(D_1(L-\tau) \leq C)\, f_{H,D}^{r-C}(\tau)\, d\tau + \mathbb{P}(D(L) \leq r - C) \geq \overline{\alpha}_1 \tag{3.15}$$

$$\mathbb{P}(D(L) \leq r - C) \geq \overline{\alpha}_2 \tag{3.16}$$

$$Q \geq 0, \tag{3.17}$$

$$r > C \geq 0, \tag{3.18}$$

where $\mathbb{E}(B_1^{\infty}(Q,r,C))$ and $\mathbb{E}(B_2^{\infty}(Q,r,C))$ are given by equations (3.9) and (3.10) respectively. We note that the constraint $r > C$ in (3.18) is implied by constraint (3.16) for demand with positive support, as the probability of that demand being less than zero equals zero and cannot be bigger or equal to $\overline{\alpha}_2 > 0$. We express this strict inequality here to remind us of what the feasible region looks like.

3.2 SLP using normal distribution as approximation of non-negative demand

A common practice in stochastic inventory models is to use the normal distribution as an approximation of the non-negative demand, i.e., the stochastic inventory models are formulated based on the characteristics of the non-negative demand and then are implemented using normal distribution as an approximation. The problem with the normal distribution is that there is always a small probability of negative demand. The normal distribution is a good approximation of non negative demand when the coefficient of variation is less than or equal to 0.5, i.e., $CV \leq 0.5$ (Peterson and Silver (1979)) in which case the probability of being less than 0 is less than 0.0228.

To solve (**SLP**) using normal distribution as approximation of the non-negative demand, the expressions that characterize the hitting time $\tau_{H,D}^{r-C}$ and τ_{H,D_1}^{C}, and the backorders, under normally distributed demand, are required. For this, consider that each class i has identical and independent normally distributed demand per unit time, with mean $\mu_i > 0$ and variance $\sigma_i^2 > 0$, $D_i(\tau) \sim N(\mu_i\tau, \sigma_i^2\tau)$, and $D(\tau) \sim N(\mu\tau, \sigma^2\tau)$, where $\mu = \mu_1 + \mu_2$ and $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Following equation (3.5), the density distribution of the hitting time $\tau_{H,D}^{r-C}$ under normally distributed demand is:

$$f_{H,D}^{r-C}(\tau) = \left(\frac{r - C + \mu\tau}{2\tau}\right) \frac{1}{\sigma\sqrt{\tau}} \varphi\left(\frac{r - C - \mu\tau}{\sigma\sqrt{\tau}}\right), \tag{3.19}$$

where $\varphi(x)$ is the density function of the standard normal distribution. In the same way, the density distribution of the hitting time τ_{H,D_1}^{C} under normally distributed demand is:

$$f_{H,D_1}^{C}(\tau) = \left(\frac{C + \mu_1\tau}{2\tau}\right) \frac{1}{\sigma_1\sqrt{\tau}} \varphi\left(\frac{C - \mu_1\tau}{\sigma_1\sqrt{\tau}}\right).$$

Then, the expected backorders in steady state given by equations (3.9) and (3.10) under normally distributed demand become:

$$\mathbb{E}(B_2^{\infty}(Q,r,C)) = \frac{\mu_2}{Q}\int_0^L \left(G\left(\frac{r - C - \mu\tau}{\sigma\sqrt{\tau}}\right) - G\left(\frac{r + Q - C - \mu\tau}{\sigma\sqrt{\tau}}\right)\right)\sigma\sqrt{\tau}d\tau\,, \tag{3.20}$$

$$\mathbb{E}(B_1^{\infty}(Q,r,C)) = \frac{\mu_1}{Q}\int_0^L \int_t^L f_{H,D_1}^{C}(t)\left(G\left(\frac{r - C - \mu(\tau - t)}{\sigma\sqrt{\tau - t}}\right) - G\left(\frac{r + Q - C - \mu(\tau - t)}{\sigma\sqrt{\tau - t}}\right)\right)\sigma\sqrt{\tau - t}\, d\tau dt \tag{3.21}$$

where $G(x) = \int_x^{\infty}(v - x)\varphi(v)dv = \varphi(x) - x(1 - \Phi(x))$ is the *loss function* (Axsäter (2007)) and $\Phi(x)$ is the distribution function of the standard normal distribution.

3.3 Solution approach

Consider the following relaxation of (**SLP**), obtained by dropping the expected backorder expressions:

Problem (**RSLP**):

$$\begin{aligned} \min_{Q,r,C} \quad & S\frac{\mu}{Q} + h\left(\frac{Q}{2} + r - \mu L\right) \\ \text{s.t:} \quad & (3.15), (3.16), (3.18). \end{aligned} \tag{3.22}$$

It is easy to show that (**RSLP**) is a relaxation of (**SLP**) because the objective function of the (**RSLP**) is less than or equal to the objective function of (**SLP**) and the feasible region is the same.

Therefore, the optimal solution of the problem (**RSLP**) is a lower bound (LB) of problem (**SLP**). Also, if we solve (**RSLP**), and then use the resulting parameters (Q, r, C) to evaluate the objective function of (**SLP**) we obtain a feasible solution and, hence, an upper bound (UB) for the problem (**SLP**). Thus, we have a method that gives a lower bound and an upper bound of the original problem.

Note that (**RSLP**) is separable in two sub-problems. The first sub problem minimizes $S\frac{\mu}{Q} + h\frac{Q}{2}$ without constraints on Q and gives the replenishment batch $Q = \sqrt{\frac{2\mu S}{h}}$ that corresponds to the EOQ problem and the second sub problem, denoted (**SLP0**), is

Problem (**SLP0**):

$$\min_{r,C} \quad r \tag{3.23}$$

$$\text{s.t:} \quad (3.15), (3.16), (3.18).$$

Therefore, the service level problem reduces to determining the optimal reorder point and critical level (r, C) that minimize the reorder point r subject to satisfying the required service levels.

To determine the optimal parameters of **(SLP0)** we take advantage of the structure of the constraints (3.15), (3.16) and (3.18). From these constraints we derive structural properties that are necessary to obtain the exact solution to the **(SLP0)** problem.

Proposition 1. *$\alpha_2(r, C)$ is increasing in r and decreasing in C and only depends on $(r - C)$.*

Proof. From equation (3.11) we have $\alpha_2(r, C) = F_{D(L)}(r - C)$. The result follows since the distribution function is a monotonically increasing function. □

The main consequence of proposition 1 is that, given a reorder point r, the maximum service level provided to the low priority class is $\alpha_2(r, 0)$.

3.3.1 Solution characterization for (SLP0): increasing non-negative demand

For any strictly increasing non-negative demand that represent the total demand of class i, with $i = 1, 2$, we obtain the following structural properties.

Proposition 2. *If $D_1(\tau)$ is a strictly increasing non-negative demand, then $\alpha_1(r, C)$ is increasing in r and C. (Proof. Appendix A)*

The main consequence of proposition 2 is that, given a reorder point r, the minimum service level provided to high priority class is $\alpha_1(r, 0)$.

Let r_i^0 be the minimum reorder point r such that the service level provided to the class i, given a critical level $C = 0$, is greater than or equal to his preset service level $\overline{\alpha}_i$, i.e., $r_i^0 = \min\{r \mid \alpha_i(r, 0) \geq \overline{\alpha}_i\}$, with $i = 1, 2$. Since functions $\alpha_i(r, 0)$ for $i = 1, 2$ are increasing in r, from propositions 1 and 2, we have that r_i^0 solves $\alpha_i(r, 0) = \overline{\alpha}_i$ for $i = 1, 2$. In particular this gives

$$r_2^0 = F_{D(L)}^{-1}(\overline{\alpha}_2). \tag{3.24}$$

Furthermore, from equation (3.13) we have that $\alpha_1(r, 0) = \alpha_2(r, 0)$ for any $r \geq 0$, because $F_{D_1(\tau)}(0) = 0$ for any $\tau > 0$. Then, since $\alpha_1(r_1^0, 0) = \overline{\alpha}_1 > \overline{\alpha}_2 = \alpha_2(r_2^0, 0) = \alpha_1(r_2^0, 0)$, and $\alpha_1(r, 0)$ is increasing in r we conclude that $0 < r_2^0 < r_1^0$ for any $\overline{\alpha}_2 > 0$.

Using proposition 1 and 2 we propose the following general solution for the **(SLP0)** problem.

Proposition 3. *If $D_i(\tau)$, with $i = 1, 2$, are strictly increasing non-negative demand and $\overline{\alpha}_2 > 0$, then the optimal parameters of the critical level policy are obtained from the equation system formed by $\alpha_1(r, C) = \overline{\alpha}_1$ and $\alpha_2(r, C) = \overline{\alpha}_2$, i.e.,*

$$r^* - C^* = F_{D(L)}^{-1}(\overline{\alpha}_2), \tag{3.25}$$

$$\int_0^L \mathbb{P}(D_1(L - \tau) \leq C^*)\, f_{H,D}^{r^*-C^*}(\tau)\, d\tau = \overline{\alpha}_1 - \overline{\alpha}_2, \tag{3.26}$$

and the service levels provided to each class are equal to their preset levels, i.e., $\alpha_i(r^*, C^*) = \overline{\alpha}_i$, $i = 1, 2$.

Proof. Let $C_2(r)$ be the maximum critical level, given a reorder point r, that ensures a service level $\overline{\alpha}_2$, i.e., $C_2(r) = \max\{C \mid \alpha_2(r, C) \geq \overline{\alpha}_2\}$. From proposition 1 we can derive that $C_2(r)$ is solution of $\alpha_2(r, C) = \overline{\alpha}_2$. Then, $C_2(r) = r - F_{D(L)}^{-1}(\overline{\alpha}_2) = r - r_2^0$, i.e., $C_2(r)$ is increasing and linear in r. Note that, $C_2(r) < r$ for any $\overline{\alpha}_2 > 0$. In the same way we define $C_1(r)$ as the minimum critical level, given a reorder point r, that ensures a service level $\overline{\alpha}_1$, i.e., $C_1(r) = \min\{C \mid \alpha_1(r, C) \geq \overline{\alpha}_1\}$. From proposition 2 we obtain that $C_1(r)$ is solution of $\alpha_1(r, C) = \overline{\alpha}_1$ and that $C_1(r)$ is strictly decreasing in r. Once $C_1(r)$ and $C_2(r)$ are defined, the feasible region of **(SLP0)** problem where all (r, C) satisfy that $\alpha_1(r, C) \geq \overline{\alpha}_1$, $\alpha_2(r, C) \geq \overline{\alpha}_2 > 0$, and $r > C \geq 0$, is the intersection of the areas above $C_1(r)$ and below $C_2(r)$. The feasible region is shown in figure 3.1.

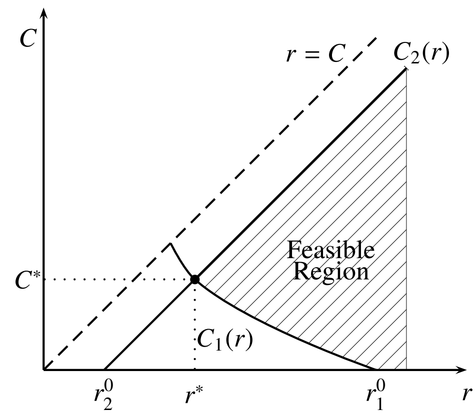


Figure 3.1: Feasible region of **SLP0** problem and $\overline{\alpha}_2 > 0$.

The figure 3.1 shows that the optimal reorder point r^* of **(SLP0)** problem occurs when $C_2(r) = C_1(r) = C^*$. Therefore, the optimal parameters of the critical level policy are obtained from the equation system formed by $\alpha_1(r, C) = \overline{\alpha}_1$ and $\alpha_2(r, C) = \overline{\alpha}_2$, and the presets service levels are satisfied exactly. Note that the existence of an r such that $C^* = C_2(r) = C_1(r)$ is guaranteed, because $0 < r_2^0 < r_1^0$ as shown above, $C_1(r)$ is strictly decreasing and continuous in r, $C_1(r_1^0) = 0$, and from equation (3.12) we obtain that there exists an $r > 0$ such that $C_1(r) = r > 0$ for any $\overline{\alpha}_1 > 0$. The argument is complete noting that $C_2(r) = r - r_2^0 < r$ is linear and increasing in r. □

Some consequences of the above proof are: (i) the optimal reorder point r^* is strictly greater than the optimal critical level C^* because $r_2^0 > 0$ when $\overline{\alpha}_2 > 0$, therefore, the constraint (3.18) may be replaced by: $r, C \geq 0$; and (ii) the optimal critical level is strictly greater than zero, i.e., $C^* > 0$, because $r_2^0 < r_1^0$.

Proposition 3 provides a general solution for (**SLP0**) when $D_i(\tau)$ of class $i = 1, 2$, are represented with strictly increasing non-negative demand, and $\overline{\alpha}_2 > 0$. Solving for the optimal solution remains challenging in general, as equations (3.25)-(3.26) have to be solved numerically and include the distribution function of $D(L)$ and the density function of $\tau_{H,D}^{r-C}$ which have to be derived from the input.

3.3.2 Solution characterization for (SLP0): normally distributed demand

Recall that we use the normal distribution as approximation of the non-negative demand. Under normally distributed demand we obtain the following structural properties.

Proposition 4. *Under normally distributed demand, the function $\alpha_1(r, C)$ is strictly increasing in r for any $0 \leq C < r$.*

Proof. Using equation (3.13), the service level provided to the high priority class using normal distribution can be write as:

$$\alpha_1(r,C) = \int_0^L \left\{ \int_{-\infty}^{\frac{C-\mu_1(L-\tau)}{\sigma_1\sqrt{L-\tau}}} \varphi(x)dx \right\} f_{H,D}^{r-C}(\tau)\, d\tau + \mathbb{P}(D(L) \le r - C)$$

and changing the order of integration we have:

$$\begin{aligned}
\alpha_1(r,C) &= \int_{\frac{C-\mu_1 L}{\sigma_1\sqrt{L}}}^{\infty} \left\{ \int_{\tau(x)}^{L} f_{H,D}^{r-C}(\tau)d\tau \right\} \varphi(x)dx + \int_{-\infty}^{\frac{C-\mu_1 L}{\sigma_1\sqrt{L}}} \left\{ \int_0^L f_{H,D}^{r-C}(\tau)d\tau \right\} \varphi(x)dx + \mathbb{P}(D(L) \le r - C) \\
&= \int_{\frac{C-\mu_1 L}{\sigma_1\sqrt{L}}}^{\infty} \{\mathbb{P}(D(\tau(x)) \le r - C) - \mathbb{P}(D(L) \le r - C)\}\, \varphi(x)dx \\
&\quad + \int_{-\infty}^{\frac{C-\mu_1 L}{\sigma_1\sqrt{L}}} \{1 - \mathbb{P}(D(L) \le r - C)\}\, \varphi(x)dx + \mathbb{P}(D(L) \le r - C) \\
&= \int_{\frac{C-\mu_1 L}{\sigma_1\sqrt{L}}}^{\infty} \mathbb{P}(D(\tau(x)) \le r - C)\varphi(x)dx + \int_{-\infty}^{\frac{C-\mu_1 L}{\sigma_1\sqrt{L}}} \varphi(x)dx \\
&= \int_{\frac{C-\mu_1 L}{\sigma_1\sqrt{L}}}^{\infty} \mathbb{P}(D(\tau(x)) \le r - C)\varphi(x)dx + \mathbb{P}(D_1(L) \le C),
\end{aligned}$$

where $\tau(x)$ is obtained from: $x\sigma_1\sqrt{L-\tau} = C - \mu_1(L-\tau)$. Although $\tau(x)$ is the result of a quadratic equation, the proof remains valid. □

Under normally distributed demand there is always a probability for negative demand. This fact makes it difficult to prove that $\alpha_1(r,C)$ is increasing in C for any $r > C \ge 0$, as we have for strictly increasing non-negative demand. We provide the expression for $\frac{\partial\alpha_1(r,C)}{\partial C}$ in (A.1) in the appendix A. Our numerical computations however have shown that such monotonicity of $\alpha_1(r,C)$ with respect to C exists for large values of the reorder point r. We therefore make this monotonicity an assumption, which we validate with computational results in section 3.4.

Assumption 1. *Assume normally distributed demand and let $\hat{r}_1$ be the solution of $\alpha_1(\hat{r}_1, 0) = 0.5$. Then, for any $r \ge \hat{r}_1$ the function $\alpha_1(r,C)$ is an increasing function of C in the interval $C \in [0,r)$.*

The main consequence of assumption 1 is that, given a reorder point $r \ge \hat{r}_1$, the minimum service level provided to high priority class is $\alpha_1(r,0)$.

From proposition 4 we derived that r_1^0 is solution of $\alpha_1(r,0) = \overline{\alpha}_1$ and under normally distributed demand we can obtain from equation (3.13) that $\alpha_1(r,0) > \alpha_2(r,0)$ for any finite $r \ge 0$. Then, as $\alpha_i(r,0)$ is increasing in r, with $i = 1,2$, we infer that the relationship between r_2^0 and r_1^0 depends on the difference $\overline{\alpha}_1 - \overline{\alpha}_2$. Thus, we have two cases: (1) $r_2^0 < r_1^0$ if $\overline{\alpha}_1 - \overline{\alpha}_2$ is large enough; or (2) $r_2^0 > r_1^0$ if $\overline{\alpha}_1 - \overline{\alpha}_2$ is small. A simple way to discriminate if we are in case 1 or 2 is to evaluate numerically $\alpha_1(r_2^0,\ 0)$. Then, if $\alpha_1(r_2^0,\ 0) < \overline{\alpha}_1$, we are in case 1, otherwise, we are in case 2. The method proposed in this chapter to solve the **(SLP0)** problem using normally distributed demand, depends on which case occurs. Note that under normally distributed demand $r_2^0 = F_{D(L)}^{-1}(\overline{\alpha}_2) = \mu L + z_{\overline{\alpha}_2}\sigma\sqrt{L} \geq \mu L > 0$ if $\overline{\alpha}_2 \geq 0.5$, where $z_{\overline{\alpha}_2}$ is the inverse standard normal distribution for a preset service level $\overline{\alpha}_2$.

Using proposition 4 and assumption 1 we propose the following solution for the **(SLP0)** problem using normally distributed demand as approximation of non-negative demand.

Proposition 5. *Under normally distributed demand, the assumption 1 and* $\overline{\alpha}_2 \in [0.5, 1)$*, the optimal parameters of the critical level policy are obtained from the following system of equations:*

(a) If $\alpha_1(r_2^0, 0) < \overline{\alpha}_1$*:*

$$r^* - C^* = \mu L + z_{\overline{\alpha}_2}\,\sigma\sqrt{L}, \tag{3.27}$$

$$\int_0^L \mathbb{P}(D_1(L-\tau) \leq C^*)\, f_{H,D}^{r^*-C^*}(\tau)\, d\tau = \overline{\alpha}_1 - \overline{\alpha}_2, \tag{3.28}$$

and the service levels provided to each class are equal to their preset levels, i.e., $\alpha_i(r^*, C^*) = \overline{\alpha}_i,\ i = 1, 2.$

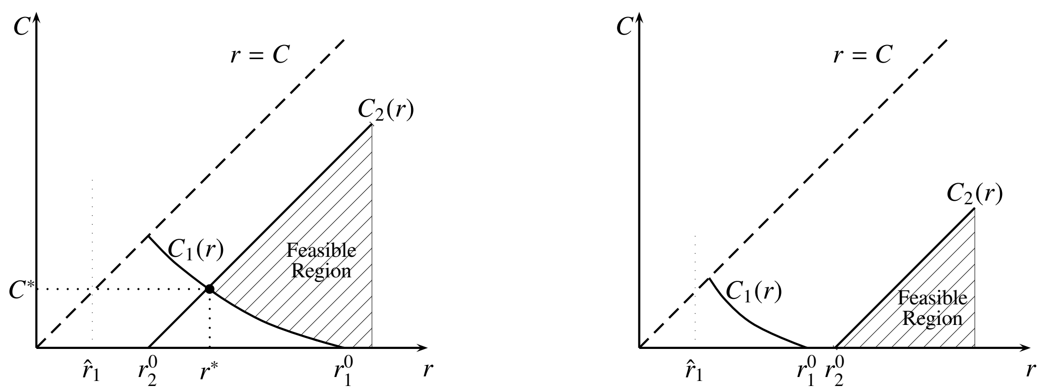
(b) If $\alpha_1(r_2^0, 0) \geq \overline{\alpha}_1$*:*

$$C^* = 0, \tag{3.29}$$

$$r^* = \mu L + z_{\overline{\alpha}_2}\,\sigma\sqrt{L}, \tag{3.30}$$

and service levels provided to each class are: $\alpha_1(r^*, 0) \geq \overline{\alpha}_1$ *and* $\alpha_2(r^*, 0) = \overline{\alpha}_2$ *for high and low priority class respectively.*

Proof. Under normally distributed demand, $C_2(r) = r - r_2^0 = r - \mu L - z_{\overline{\alpha}_2}\,\sigma\sqrt{L}$, and continues to be increasing and linear in r. On the other hand, from proposition 4 we can obtain that $C_1(r)$ is solution of $\alpha_1(r, C) = \overline{\alpha}_1$ and under assumption 1 we can conclude that $C_1(r)$ is strictly decreasing in r at least from some $r > C \geq \hat{r}_1$ until $r \leq r_1^0$. Then, under normally distributed demand, the feasible region of **(SLP0)** problem using normally distributed demand where all (r, C) satisfy that $\alpha_1(r, C) \geq \overline{\alpha}_1$, $\alpha_2(r, C) \geq \overline{\alpha}_2$, with $\overline{\alpha}_2 \in [0.5, 1)$, and $r > C \geq 0$, is the same as defined for the proof of proposition 3, but in this case, it can take two different forms, shown in figure 3.2. If $\alpha_1(r_2^0, 0) < \overline{\alpha}_1$, then $r_2^0 < r_1^0$ which induces the first feasible regions shown in figure (3.2a) when $\overline{\alpha}_2 \in [0.5, 1)$. If $\alpha_1(r_2^0, 0) \geq \overline{\alpha}_1$, then $r_2^0 > r_1^0$, which induces a second feasible region shown in figure (3.2b) when $\overline{\alpha}_2 \in [0.5, 1)$. Note that $r_2^0 \geq \mu L > \hat{r}_1$ when $\overline{\alpha}_2 \in [0.5, 1)$.



(a) $\alpha_1(r_2^0,\ 0) < \overline{\alpha}_1$ and $\overline{\alpha}_2 \in [0.5, 1)$ **(b)** $\alpha_1(r_2^0,\ 0) \geq \overline{\alpha}_1$ and $\overline{\alpha}_2 \in [0.5, 1)$

Figure 3.2: Feasible regions for **SLP0** problem using normally distributed demand and $\overline{\alpha}_2 \in [0.5, 1)$

The figure (3.2a) shows that the optimal reorder point r^* of the **(SLP0)** problem using normally distributed demand occurs when $C_2(r) = C_1(r) = C^*$. Therefore, the optimal parameters of the critical level policy are obtained from the equation system formed by $\alpha_1(r, C) = \overline{\alpha}_1$ and $\alpha_2(r, C) = \overline{\alpha}_2$. From figure (3.2b) we conclude that the minimum reorder point that guarantees a service level $\overline{\alpha}_1$ provided to high priority class and a service level $\overline{\alpha}_2$ provided to the low priority class is r_2^0. Therefore, $r^* = r_2^0 = \mu L + z_{\overline{\alpha}_2}\sigma\sqrt{L}$ and $C^* = 0$. □

Some consequences of the above proof are: (i) given equation (3.27), the equation (3.28) only depends on C^*; (ii) if $\alpha_1(r_2^0, 0) < \overline{\alpha}_1$, then r_2^0 is a lower bound of **(SLP0)** problem using normally distributed demand when $\overline{\alpha}_2 \in [0.5, 1)$; (iii) if $\alpha_1(r_2^0, 0) < \overline{\alpha}_1$ and $\overline{\alpha}_2 \in [0.5, 1)$, then $r^* > C^*$, because $r_2^0 \geq \mu L > 0$, therefore, constraint (3.18) may be replaced by: $r, C \geq 0$; and (iv) if $\alpha_1(r_2^0, 0) < \overline{\alpha}_1$ and $\overline{\alpha}_2 \in [0.5, 1)$, then $C^* > 0$, because $0 < r_2^0 < r_1^0$ when $\overline{\alpha}_2 \in [0.5, 1)$.

The proposition 5 is a general solution for the **(SLP0)** problem using normally distributed demand under assumption 1 and $\overline{\alpha}_2 \in [0.5, 1)$. On the other hand, similar to the general solution in the case with demands with non-negative support, it can be difficult to compute the critical level C^* from equation (3.28).

The following results compare the reorder point induced by the critical level policy with the reorder point induced by the round-up policy and separate stock policy. Let r_u be the reorder point induce by the round-up policy and r_s be the reorder point induced by the separate stock policy. The reorder point of the round-up policy is obtained from $F_{D(L)}(r_u) = \overline{\alpha}_1$ and the reorder point of the separate-stock policy is obtained from $r_s = r_1^s + r_2^s$, where r_1^s is solution of $F_{D_1(L)}(r_1^s) = \overline{\alpha}_1$ and r_2^s is solution of $F_{D_2(L)}(r_2^s) = \overline{\alpha}_2$. Under normally distributed demand, $r_u = \mu L + z_{\overline{\alpha}_1}\sigma\sqrt{L}$, and $r_s = \mu L + z_{\overline{\alpha}_1}\sigma_1\sqrt{L} + z_{\overline{\alpha}_2}\sigma_2\sqrt{L}$, where $z_{\overline{\alpha}_1}$ is the inverse standard normal distribution for a preset service level $\overline{\alpha}_1$. Note that, under normal distributed demand, $r_u \leq r_s$ if $z_{\overline{\alpha}_1} \leq z_{\overline{\alpha}_2}\frac{\sigma_2}{\sigma-\sigma_1}$ and that $\frac{\sigma_2}{\sigma-\sigma_1} > 1$.

Proposition 6. *Under normally distributed demand, the assumption 1 and* $\overline{\alpha}_2 \in [0.5, 1)$*, the optimal reorder point of the critical level policy is strictly less than the reorder point induced by the round-up policy, i.e.,* $r^* < r_u$*, and strictly less than the reorder point induced by the separate stock policy, i.e.,* $r^* < r_s$*, when* $\alpha_1(r_2^0,\ 0) \geq \overline{\alpha}_1$*,*

Proof. From equation (3.13) we note that $\alpha_1(r_u, 0) > \overline{\alpha}_1$ because we assume that the lead time and the parameters of the demand per unit time of both classes are finite and $\overline{\alpha}_1 < 1$. From propositions 5(a) we obtained that $\alpha_1(r^*, C^*) = \overline{\alpha}_1$ and $r^* > C^*$, and from assumption 1 we derive that $\alpha_1(r^*, C^*) = \overline{\alpha}_1 \geq \alpha_1(r^*, 0)$. Therefore, it holds that $\alpha_1(r_u, 0) > \overline{\alpha}_1 \geq \alpha_1(r^*, 0)$ and from proposition 4 we conclude that $r_u > r^*$. On the other hand, from proposition 5(b) we obtained that $\alpha_2(r^*, 0) = F_{D(L)}(r^*) = \overline{\alpha}_2$. By definition, $\overline{\alpha}_1 > \overline{\alpha}_2$, then $F_{D(L)}(r_u) > F_{D(L)}(r^*)$, and we conclude that $r_u > r^*$.

Following similar logic to compare the optimal reorder point of the critical level policy with respect to the reorder point induced by the separate stock policy, from proposition 5(b) we obtain that $r^* = \mu L + z_{\overline{\alpha}_2}\sigma\sqrt{L}$. Since $\overline{\alpha}_1 > \overline{\alpha}_2$, we conclude that $r^* < r_s$ from the triangle inequality. □

Unfortunately, we have not found a simple proof that the reorder point induced by the critical level policy is strictly less than the reorder point induced by separate stock policy when $\alpha_1(r_2^0,\ 0) < \overline{\alpha}_1$ and $\overline{\alpha}_2 \in [0.5, 1)$. The derivation requires checking that

$$\int_0^L \mathbb{P}\left(D_1(L-\tau) \leq z_{\overline{\alpha}_1}\sigma_1\sqrt{L} + z_{\overline{\alpha}_2}\sigma_2\sqrt{L} - z_{\overline{\alpha}_2}\sigma\sqrt{L}\right) f_{H,D}^{r_2^0-0}(\tau)\,\partial\tau > \overline{\alpha}_1 - \overline{\alpha}_2\,, \tag{3.31}$$

which is not much different from solving the system of equations (3.27)-(3.28) and seeing if $r^* < r_s$.

3.4 Computational study

In this section, we present our numerical study and its results. The main objective of the computational study is to show how good is the performance of our solution approach, compared the critical level policy with the separate stock and round-up policies and provide numerical evidence to validate assumption 1.

For simplicity, we use normally distributed demand as an approximation to the non-negative demand, solving **RSLP**, from which we obtain a lower bound of **SLP**. Let (Q^*, r^*, C^*) be the optimal critical level policy controls of **RSLP**; $AC_{RSLP}(Q^*, r^*, C^*)$ be the objective function of **RSLP**; and $AC_{SLP}(Q^*, r^*, C^*)$ be the objective function of **SLP** given the optimal critical level policy controls of **RSLP**. Note that $AC_{RSLP}(Q^*, r^*, C^*) = LB$, $AC_{SLP}(Q^*, r^*, C^*) = UB$ and $UB > LB$. In order to evaluate the performance of our solution, we carried out several test problems and computed the percentage of optimality gap, $Gap(\%)$, expressed as $100 \times (UB - LB)/LB$.

Recall that **RSLP** is separable in the **EOQ** and **SLP0** problems. Then, for each test problem we determine the replenishment batch solving the EOQ model, i.e., $Q^* = \sqrt{2\mu S/h}$, and the reorder point r^* and the optimal critical level C^* solving the system of non-linear equations given in proposition 5.

The equation systems of proposition 5 were programmed by a C code using Brent-Dekker method. Backorders in the steady state given by equations (3.20) and (3.21) were also programmed in C code, like the numerical experiments to validate assumption 1. All test were carried on a PC with Intel Core i7 2.3 GHz processor and 16 GB RAM. The time to compute the parameters of the critical level policy are on average 0.0011 seconds and in the worst case 0.0019 seconds.

3.4.1 Experimental result for (SLP) problem using normally distributed demand

In order to cover a wide range of data, we design a set of 10 experiments to evaluate the performance of our solution approach and to compare the critical level policy with the separate stock and round-up policies. In each experiment we fix the preset service levels $\overline{\alpha}_1$ and $\overline{\alpha}_2$, and consider a base case with the following parameters: normal demand distributions with mean $\mu_1 = \mu_2 = 25$ and coefficient of variation $CV_1 = CV_2 = 0.2$ ($\sigma_1^2 = \sigma_2^2 = 25$), lead time $L = 5$, ordering cost $S = 300$ and holding cost per unit and unit time $h = 0.75$. We conduct experiments studying the sensitivity of the solutions to changing parameters CV_i, μ_i, S, and h. This gives a total of 135 experiments for each setting of the preset service levels.

Our numerical results show that our solution approach is able to provide good-quality solutions that are on average 0.3% and at worst 7.8% from the optimal solution. Table 3.1 show the average and maximum relative gap over 45 instances for the ten settings of preset service levels and different values of S.

		Gap(%)					
		$S = 100$		$S = 300$		$S = 500$	
$\overline{\alpha}_1$	$\overline{\alpha}_2$	Average	Max	Average	Max	Average	Max
0.975	0.55	1.37	7.81	0.55	3.36	0.35	2.15
0.975	0.65	0.81	4.39	0.33	1.95	0.21	1.26
0.975	0.75	0.44	2.26	0.18	1.02	0.12	0.67
0.975	0.85	0.19	0.93	0.08	0.43	0.05	0.28
0.975	0.95	0.05	0.18	0.02	0.09	0.01	0.06
0.800	0.75	0.87	3.03	0.34	1.32	0.22	0.85
0.850	0.75	0.70	2.80	0.28	1.22	0.18	0.79
0.900	0.75	0.58	2.59	0.23	1.14	0.15	0.74
0.950	0.75	0.48	2.38	0.20	1.06	0.13	0.69
0.999	0.75	0.38	2.04	0.16	0.93	0.10	0.61

Table 3.1: Optimality Gap(%) between lower and upper bounds

Table 3.1 shows that the maximum relative gap occurs when there is maximum difference between the preset service levels and the ordering cost is minimal ($S = 100$). As an example, table 3.2 shows the relative optimality gap for the 135 problems of the experiment: $\overline{\alpha}_1 = 0.975$ and $\overline{\alpha}_2 = 0.75$.

			Gap(%)								
			$\mu_1 = 100, \mu_2 = 25$			$\mu_1 = \mu_2 = 25$			$\mu_1 = 25, \mu_2 = 100$		
CV_1	CV_2	h	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$
0.2	0.2	0.25	0.03	0.01	0.01	0.02	0.01	0.00	0.11	0.04	0.02
		0.75	0.07	0.03	0.02	0.06	0.02	0.01	0.29	0.11	0.06
		1.25	0.11	0.04	0.03	0.09	0.03	0.02	0.46	0.17	0.11
0.4	0.4	0.25	0.09	0.03	0.02	0.07	0.03	0.02	0.35	0.13	0.08
		0.75	0.21	0.09	0.06	0.19	0.07	0.04	0.91	0.35	0.22
		1.25	0.30	0.13	0.09	0.29	0.11	0.07	1.36	0.55	0.35
0.6	0.6	0.25	0.16	0.06	0.04	0.14	0.05	0.03	0.67	0.25	0.16
		0.75	0.35	0.16	0.10	0.36	0.14	0.09	1.60	0.67	0.43
		1.25	0.48	0.23	0.16	0.53	0.22	0.14	2.26	1.02	0.67
0.6	0.2	0.25	0.15	0.06	0.04	0.09	0.03	0.02	0.15	0.05	0.03
		0.75	0.34	0.15	0.10	0.22	0.09	0.05	0.40	0.15	0.09
		1.25	0.46	0.22	0.15	0.34	0.14	0.09	0.62	0.24	0.15
0.2	0.6	0.25	0.04	0.01	0.01	0.09	0.03	0.02	0.65	0.24	0.15
		0.75	0.10	0.04	0.02	0.23	0.09	0.05	1.57	0.65	0.41
		1.25	0.15	0.06	0.04	0.35	0.14	0.09	2.23	0.99	0.65

Table 3.2: Optimality Gap(%) when $\overline{\alpha}_1 = 0.975$ and $\overline{\alpha}_2 = 0.75$

The pattern of behavior of the relative optimality gap observed in table 3.2 is repeated for all ten experiments, i.e., the relative gap is decreasing in S and increasing in h. The maximum gap occurs when class 2 dominates on mean and variance ($\mu_2 = 100$ and $CV_2 = 0.6$), the ordering cost is minimal ($S = 100$) and the holding cost per unit and unit time is maximum ($h = 1.25$). Note that we obtain Q from EOQ problem, therefore, for a low ordering cost and high holding cost per unit and unit time, we hope a low batch size. Then, for a low batch size and domain of low priority class in mean and variance, we expected high backorders class 2 for the critical level policy. Since our solution approach is based on a relaxation which despises backorders of class 2, we expect a high gap between lower and upper bound when the parameters induce high backorders of class 2.

In the next set of results we compare the efficiency of the critical level policy obtained with the proposed approach against the separate stock and round-up policies. For every one of the 135 problem parameters considered above and ten preset service levels settings, we determine the three different policies and compute for each the operational costs. As we expected, the critical level policy outperformed both the separate stock and round-up policies in the 1350 problems considered. Table 3.3 shows the average and maximum relative benefit of the critical level policy with respect to the round-up and separate stock for the 10 settings of preset service levels and different values of S.

Table 3.3 shows that in all experiments, the average relative benefit is greater with respect to the separate stock policy, but the maximum relative benefit is reached when comparing against the round-up policy. We also note that the relative benefit to the round-up is more sensitive and, by contrast, using two separate lot sizes and two separate reorder points causes a more homogeneous benefit. The maximum relative benefit, with respect to round-up or separate stock policies, occurs when there is maximum difference between the preset service levels and the ordering cost is minimal ($S = 100$). As an example, table 3.4 shows the relative benefit regarding round-up and separate stock for the 135 problems of the experiment: $\overline{\alpha}_1 = 0.975$ and $\overline{\alpha}_2 = 0.75$.

The pattern of the maximum relative benefit regarding round-up policy, observed in table 3.4, is repeated for all ten experiments, i.e., the maximum benefit occurs when the class 2 dominates on mean and variance ($\mu_2 = 100$, $CV_2 = 0.6$), the ordering cost is minimal ($S = 100$) and the holding cost per unit and unit time is maximum ($h = 1.25$). Clearly, the round-up policy is highly inefficient when the class 2 dominates mean and variance, because under this situation, this policy provides too much inventory to the low priority class causing a high reorder point and therefore a high cost. On the other hand, when ordering cost is low and holding cost per unit and unit time is high, batch

		Benefit (%) vs Round-up					
		$S = 100$		$S = 300$		$S = 500$	
$\overline{\alpha}_1$	$\overline{\alpha}_2$	Average	Max	Average	Max	Average	Max
0.975	0.55	18.27	46.88	13.45	38.43	11.34	34.07
0.975	0.65	16.01	41.10	11.73	33.44	9.87	29.57
0.975	0.75	13.38	34.16	9.76	27.62	8.20	24.36
0.975	0.85	9.96	24.95	7.24	20.06	6.08	17.65
0.975	0.95	4.04	8.97	2.93	7.19	2.46	6.33
0.800	0.75	3.74	8.08	2.52	5.90	2.04	4.96
0.850	0.75	5.89	14.07	4.04	10.63	3.30	9.07
0.900	0.75	8.24	20.45	5.77	15.84	4.76	13.68
0.950	0.75	11.21	28.41	8.04	22.53	6.70	19.68
0.999	0.75	19.63	48.15	15.02	40.94	13.03	37.06

		Benefit (%) vs Separate stock					
		$S = 100$		$S = 300$		$S = 500$	
$\overline{\alpha}_1$	$\overline{\alpha}_2$	Average	Max	Average	Max	Average	Max
0.975	0.55	24.56	32.43	25.27	31.08	25.53	30.79
0.975	0.65	24.55	31.65	25.24	30.62	25.51	30.40
0.975	0.75	24.55	30.79	25.22	30.09	25.49	29.96
0.975	0.85	24.56	29.77	25.21	29.43	25.47	29.41
0.975	0.95	24.61	30.10	25.20	29.88	25.45	29.78
0.800	0.75	26.28	31.21	26.43	30.52	26.49	30.28
0.850	0.75	25.99	30.82	26.23	30.29	26.32	30.09
0.900	0.75	25.59	30.38	25.96	30.04	26.10	29.91
0.950	0.75	25.02	30.45	25.56	30.11	25.77	29.97
0.999	0.75	22.97	32.14	24.05	30.18	23.67	24.67

Table 3.3: Benefit of the critical level vs. Round-up and Separate stock policies

sizes are small and the expected backorder increases. We observe that the expected backorders induced by the critical level are greater than those induced by the round-up policy, but its effect on cost is relatively low compared with the effect of the reorder point. Note that, as Deshpande et al. (2003) observed, the relative benefit regarding Round-up is decreasing in S.

Finally, we analyze how the preset service levels impact the total cost. From equations (3.27) and (3.28) we conclude that increasing the preset service level of the high priority class causes the optimal critical level C^* and reorder point r^* to increase. Consequently, we expect an increase in the holding and total costs. In the same way, we conclude from equations (3.27) and (3.28) that increasing the preset service level of the low priority class causes the optimal reorder point to increase and the optimal critical level to decrease. Therefor, we expect an increase in the holding cost per unit time, but smaller than when $\overline{\alpha}_1$ increases. Table 3.5 shows how $AC_{SLP}(Q^*, r^*, C^*)$ increases, for different input parameters, when $\overline{\alpha}_2 = 0.75$ and the preset service level of the high priority class increase from $\overline{\alpha}_1 = 0.975$ to $\overline{\alpha}_1 = 0.999$.

From table 3.5 we observe that increase the service level of the high priority class from $\overline{\alpha}_1 =$

			Benefit(%) vs Round-up								
			$\mu_1 = 100, \mu_2 = 25$			$\mu_1 = \mu_2 = 25$			$\mu_1 = 25, \mu_2 = 100$		
CV_1	CV_2	h	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$
0.2	0.2	0.25	2.97	1.90	1.52	4.33	2.65	2.10	11.59	7.41	5.93
		0.75	4.40	2.97	2.43	6.80	4.33	3.46	17.17	11.59	9.47
		1.25	5.16	3.59	2.97	8.26	5.37	4.33	20.14	14.02	11.59
0.4	0.4	0.25	4.93	3.38	2.78	7.53	4.84	3.89	18.87	12.95	10.64
		0.75	6.67	4.93	4.17	11.04	7.53	6.17	25.54	18.87	15.97
		1.25	7.48	5.73	4.93	12.89	9.07	7.53	28.60	21.93	18.87
0.6	0.6	0.25	6.32	4.57	3.84	9.98	6.69	5.45	23.83	17.24	14.47
		0.75	8.09	6.32	5.49	13.91	9.98	8.35	30.43	23.83	20.70
		1.25	8.83	7.16	6.32	15.80	11.75	9.98	33.21	26.97	23.83
0.6	0.2	0.25	6.04	4.35	3.65	7.36	4.79	3.86	12.48	8.12	6.54
		0.75	7.75	6.04	5.24	10.65	7.36	6.07	18.06	12.48	10.28
		1.25	8.48	6.85	6.04	12.34	8.82	7.36	20.91	14.95	12.48
0.2	0.6	0.25	3.76	2.44	1.97	9.15	5.95	4.79	24.36	17.55	14.71
		0.75	5.45	3.76	3.10	13.25	9.15	7.53	31.24	24.36	21.12
		1.25	6.31	4.50	3.76	15.37	10.96	9.15	34.16	27.62	24.36

			Benefit(%) vs Separate stock								
			$\mu_1 = 100, \mu_2 = 25$			$\mu_1 = \mu_2 = 25$			$\mu_1 = 25, \mu_2 = 100$		
CV_1	CV_2	h	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$
0.2	0.2	0.25	24.06	24.59	24.77	29.25	29.27	29.27	24.62	24.95	25.06
		0.75	23.32	24.06	24.33	29.23	29.25	29.26	24.12	24.62	24.79
		1.25	22.89	23.75	24.06	29.21	29.24	29.25	23.81	24.41	24.62
0.4	0.4	0.25	23.17	23.94	24.23	29.16	29.21	29.23	23.99	24.53	24.71
		0.75	22.20	23.17	23.56	29.10	29.16	29.19	23.28	23.99	24.26
		1.25	21.71	22.74	23.17	29.07	29.14	29.16	22.90	23.68	23.99
0.6	0.6	0.25	22.56	23.45	23.80	29.06	29.14	29.17	23.54	24.18	24.42
		0.75	21.56	22.56	23.00	28.98	29.06	29.10	22.79	23.54	23.86
		1.25	21.10	22.10	22.56	28.94	29.02	29.06	22.42	23.20	23.54
0.6	0.2	0.25	21.56	22.79	23.26	29.96	29.73	29.64	28.43	27.33	26.95
		0.75	20.12	21.56	22.16	30.25	29.96	29.84	29.95	28.43	27.87
		1.25	19.43	20.90	21.56	30.39	30.09	29.96	30.79	29.09	28.43
0.2	0.6	0.25	23.42	24.16	24.42	25.40	26.89	27.39	18.93	21.27	22.10
		0.75	22.46	23.42	23.79	23.25	25.40	26.16	15.88	18.93	20.11
		1.25	21.96	23.00	23.42	22.03	24.48	25.40	14.27	17.59	18.93

Table 3.4: Benefit(%) vs. Round-up and Separate stock when $\overline{\alpha}_1 = 0.975$ and $\overline{\alpha}_2 = 0.75$

0.975 to $\overline{\alpha}_1 = 0.999$, increases the total cost in 9.8% average, and as expected, the maximum increase of $AC_{SLP}(Q^*, r^*, C^*)$ occurs when the class 1 dominates on mean and variance ($\mu_1 = 100$ and $CV_1 = 0.6$), the ordering cost is minimal ($S = 100$) and the holding cost per unit and unit time is maximum ($h = 1.25$). This is because, when class 1 is larger in mean and variance, more items are reserved for the high priority class and the threshold level C increases. On the other hand, a low ordering cost and high holding cost per unit and unit time produce a small batch size and high backorders for class 2. Then, the high class 2 backorders cause the holding cost per unit time to increase. The result is a higher holding cost and thus a higher total cost. Table 3.6 highlights how $AC_{SLP}(Q^*, r^*, C^*)$ increases for different imput parameters, when $\overline{\alpha}_1 = 0.975$ and the preset service level of low priority class increase from $\overline{\alpha}_2 = 0.75$ to $\overline{\alpha}_2 = 0.85$.

From table 3.6 we observe that increase the service level of the low priority class from $\overline{\alpha}_2 = 0.75$

			Increase (%) the total cost								
			$\mu_1 = 100, \mu_2 = 25$			$\mu_1 = \mu_2 = 25$			$\mu_1 = 25, \mu_2 = 100$		
CV_1	CV_2	h	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$
0.2	0.2	0.25	10.06	6.59	5.32	4.28	2.62	2.07	2.98	1.83	1.45
		0.75	14.47	10.06	8.32	6.75	4.28	3.42	4.65	2.98	2.39
		1.25	16.72	12.01	10.06	8.20	5.32	4.28	5.62	3.69	2.98
0.4	0.4	0.25	16.22	11.51	9.59	7.96	5.12	4.11	5.37	3.48	2.80
		0.75	21.24	16.22	13.95	11.71	7.96	6.52	7.79	5.37	4.42
		1.25	23.47	18.56	16.22	13.70	9.60	7.96	9.04	6.44	5.37
0.6	0.6	0.25	20.36	15.26	13.02	11.03	7.39	6.02	7.29	4.94	4.04
		0.75	25.21	20.36	17.98	15.41	11.03	9.22	10.02	7.29	6.13
		1.25	27.19	22.69	20.36	17.54	13.00	11.03	11.30	8.53	7.29
0.6	0.2	0.25	20.31	15.19	12.94	10.13	6.64	5.36	5.53	3.49	2.78
		0.75	25.21	20.31	17.91	14.55	10.13	8.38	8.35	5.53	4.48
		1.25	27.22	22.66	20.31	16.80	12.09	10.13	9.89	6.74	5.53
0.2	0.6	0.25	11.00	7.32	5.95	6.57	4.22	3.38	4.77	3.18	2.58
		0.75	15.48	11.00	9.17	9.70	6.57	5.38	6.67	4.77	3.98
		1.25	17.71	13.01	11.00	11.35	7.94	6.57	7.59	5.62	4.77

Table 3.5: $AC_{SLP}(Q^*, r^*, C^*)$ increase (%) between using $\bar{\alpha}_1 = 0.999$ and $\bar{\alpha}_1 = 0.975$.

			Increase (%) the total cost								
			$\mu_1 = 100, \mu_2 = 25$			$\mu_1 = \mu_2 = 25$			$\mu_1 = 25, \mu_2 = 100$		
CV_1	CV_2	h	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$
0.2	0.2	0.25	0.73	0.46	0.37	1.20	0.73	0.57	3.50	2.17	1.72
		0.75	1.08	0.73	0.59	1.92	1.20	0.95	5.40	3.50	2.82
		1.25	1.27	0.88	0.73	2.35	1.50	1.20	6.47	4.31	3.50
0.4	0.4	0.25	1.07	0.73	0.60	2.06	1.30	1.04	5.97	3.93	3.17
		0.75	1.44	1.07	0.90	3.09	2.06	1.68	8.47	5.97	4.95
		1.25	1.61	1.24	1.07	3.64	2.51	2.06	9.69	7.09	5.97
0.6	0.6	0.25	1.23	0.89	0.74	2.70	1.77	1.43	7.78	5.37	4.41
		0.75	1.55	1.23	1.07	3.85	2.70	2.24	10.41	7.78	6.61
		1.25	1.68	1.38	1.23	4.41	3.21	2.70	11.58	9.00	7.78
0.6	0.2	0.25	1.14	0.82	0.69	1.88	1.21	0.97	3.67	2.32	1.85
		0.75	1.44	1.14	0.99	2.77	1.88	1.54	5.52	3.67	2.98
		1.25	1.56	1.29	1.14	3.24	2.27	1.88	6.52	4.47	3.67
0.2	0.6	0.25	0.96	0.62	0.50	2.69	1.71	1.37	8.15	5.58	4.58
		0.75	1.40	0.96	0.79	4.00	2.69	2.19	10.99	8.15	6.90
		1.25	1.62	1.15	0.96	4.70	3.26	2.69	12.27	9.46	8.15

Table 3.6: $AC_{SLP}(Q^*, r^*, C^*)$ increase (%) between using $\overline{\alpha}_2 = 0.85$ and $\overline{\alpha}_2 = 0.75$

to $\overline{\alpha}_2 = 0.85$, increases the total cost in 3% average, and as expected, the maximum increase of $AC_{SLP}(Q^*, r^*, C^*)$ occurs when the class 2 dominates on mean and variance ($\mu_2 = 100$ and $CV_2 = 0.6$), the ordering cost is minimal ($S = 100$) and the holding cost per unit and unit time is maximum ($h = 1.25$). This is because, high class 2 backorders are produced in the critical level policy when class 2 dominates and the batch size is small.

The results obtained by using normally distributed demand are similar to those obtained by using a Poisson process (obtained from Deshpande et al. (2003)). However, using normally distributed demand as an approximation of strictly increasing non-negative demand allows us to observe the effect of changes in variance on the critical level policy, i.e., how changing the coefficient of variation CV (ratio of standard deviation to mean) affects our results. For any parameters setting that we tested, we observe that increasing σ_i^2, equivalent to increasing CV_i for a fixed μ_i, causes an increase in $AC_{SLP}(Q^*, r^*, C^*)$. This is because, we expect high backorders and large reorder point r and critical level C, when variance of class i increases for $i = 1$ or 2. Furthermore, we observe from table 3.4 that:

- benefit with respect to round-up is increasing in CV_i for $i = 1, 2$,
- benefit with respect to separate stock is decreasing in CV_1 when μ_1 dominates and increasing when $\mu_1 = \mu_2$ or μ_2 dominates,
- benefit with respect to separate stock is decreasing in CV_2.

3.4.2 Numerical evidence: assumption 1

Proposition 5 that characterizes the optimal critical level solution in the case of normally distributed demand makes the assumption that $\alpha_1(r, C)$ is increasing in $C \in [0, r)$ for a large enough r (assumption 1). In order to numerically validate this assumption and cover a wide range of data, a set of 8 experiments were designed. In each experiment we generate 100000 random sets of $\{L, \mu_1, CV_1, \mu_2, CV_2\}$, within predefined limits that appear in table 3.7. For each randomly generated set of parameters, we also generate a random reorder point in the interval $[\mu L, \mu L + z_{0.9999}\sigma\sqrt{L}]$ and a random critical level in the interval $[0, r)$. Then, for each random set $\{L, \mu_1, CV_1, \mu_2, CV_2, r, C\}$ we evaluate $\frac{\partial \alpha_1(r,C)}{\partial C}$ and $\alpha_1(r, 0)$. We provide the expression for $\frac{\partial \alpha_1(r,C)}{\partial C}$ in (A.1) in the appendix A. For each experiment we obtain the minimum $\alpha_1(r, 0)$ such that the service level provided to the high priority class is increasing in C, i.e., $\min\{\alpha_1(r, 0) \mid \partial\alpha_1(r, C)/\partial C \geq 0\}$. The first experiment randomly vary the parameters within the limits of the base case. Then, the limits of these parameters are varied. Table 3.7 shows the parameters limits at each experiment and the result obtained.

Based on the results shown in Table 3.7, we infer that for any $r \geq \hat{r}_1$, with $\hat{r}_1$ solution of $\alpha_1(\hat{r}_1, 0) = 0.5$, the function $\alpha_1(r, C)$ is an increasing function of C in the interval $C \in [0, r)$, i.e., we infer that assumption 1 seems to be valid at least within the limits of the experiments of Table 3.7.

Exp.	L	μ_1	CV_1	μ_2	CV_2	$\{\min\{\alpha_1(r,0) \mid \partial\alpha_1(r,C)/\partial C \geq 0\}\}$
1	[1,5]	(0,25]	(0,0.2]	(0,25]	(0,0.2]	0.4268
2	[1,25]	**(25,100]**	(0,0.2]	(0,25]	(0,0.2]	0.4206
3	[1,25]	(0,25]	**(0.2,2]**	(0,25]	(0,0.2]	0.3334
4	[1,25]	(0,25]	(0,0.2]	**(25,100]**	(0,0.2]	0.3715
5	[1,25]	(0,25]	(0,0.2]	(0,25]	**(0.2,2]**	0.4813
6	[1,25]	**(25,100]**	(0,0.2]	**(25,100]**	(0,0.2]	0.3642
7	[1,25]	(0,25]	**(0.2,2]**	(0,25]	**(0.2,2]**	0.4346
8	[1,25]	**(25,100]**	**(0.2,2]**	**(25,100]**	**(0.2,2]**	0.4535

Table 3.7: Numerical evidence for assumption 1

3.5 Conclusions

In this chapter we analyzed the constant critical level policy for fast-moving items when rationing is used to provide differentiated service levels to two demand classes (high and low priority). The inventory system operates under continuous review (Q, r) policy, with a critical threshold value C, full-backorder, deterministic lead time, and the service level provided to each class is measured by service level type 1.

Using the hitting time approach and the threshold clearing mechanism to satisfy backorders when multiple outstanding orders exist, we develop expressions for service levels type I under rationing and expected backorders of high and low priority classes. We formulate a service level problem as a nonlinear problem with chance constraints (service level constraints) to determine the optimal parameters of the critical level policy. We propose to optimally solve a relaxation, which allows us to obtain good-quality bounds. For strictly increasing non-negative demand, we characterize the optimal solution of the relaxed service level problem through a system of equations and, under mild assumptions, when normally distributed demand is used as approximation of the non-negative demand.

The computational results show that our solution approach can find good-quality solutions that are on average 0.3% and at worst 7.8% from the optimal solution. Given the nature of our relaxation, the maximum gap(%) occurs when the class 2 dominates on mean and variance, the ordering cost is minimal, the holding cost per unit and unit time is maximum and difference between the preset service levels is maximum.

As expected, the critical level policy outperformed both the separate stock and round-up policies in terms of total cost. Using normally distributed demand as an approximation of strictly increasing non-negative demand allows us to observe the effect of varying the coefficient of variation of the demand distribution, situation that is not possible with Poisson demand process. We observe that the benefit of critical level policy with respect round-up is increasing in the variance of the demand distributions for both classes, and the benefit with respect separate stock is decreasing in the variance of class 2 demand, decreasing in the variance of class 1 demand when the mean of class 1 demand is larger and increasing when the mean of class 2 dominates. In addition, we observe the following managerial insights:

- *the average savings induced by the critical level policy are greater with respect to separate stock, but the maximum savings are achieved when comparing to round up policy.*
- *critical level policy leads to significant savings with respect to round-up when class 2 dominates on mean and variance, the ordering cost is minimal, holding cost per unit and unit time is maximum and difference between preset service levels is maximum.*
- *critical level policy leads to significant savings with respect to separate stock when class 2 dominates on mean, class 1 dominates on variance, the ordering cost is minimal, holding cost per unit and unit time is maximum and difference between preset service levels is maximum.*
- *the cost of increasing the service level of the high priority class is significantly greater than the cost of increasing the service level of the low priority class.*

There are a number of questions and issues left for future research. The first one, is to solve exactly the SLP problem or solve a relaxation that does not drop backorders. Second is to expand the results to more than two classes. Third is to broaden the measures of service level. For instance we could use the fill-rate as service level measures, leading to different problems and therefore different solutions. In particular, the fill-rate or ready rate depend of the replenishment batch quantity, therefore, although we consider a relaxation, the service level problem is not separable as in our case. Therefore, the problem becomes more difficult to solve because the replenishment batch quantity, the reorder point and the critical level must be optimized jointly in the same service level problem. Another line of future work is to propose a cost optimization problem where backorders of each class are penalized with different cost.

4 | Convex approximation for backorders of a rationing inventory policy with fast moving items

This chapter analyzes the constant critical level policy for fast-moving items of an inventory system facing random demands from two customer classes (high and low priority). We consider a continuous review (Q, r, C) policy with continuously distributed demands. Using the *state-dependent demand* approach and the *threshold clearing mechanism* we obtain an approximation for on-hand inventory based on a convex approximation of backorders. We propose a cost optimization problem that uses the convex objective function to determine the parameters of the critical level policy. For instances we tested, computational results show that the proposed approximation allows us to obtain good-quality solutions that induce a benefit on average 7.5% and 26.3% against the round-up and separate stock policies respectively.

4.1 Model Framework

Consider a facility that holds inventory of a single type of product to serve two demand classes $i = 1, 2$, where class 1 is high priority and class 2 is low priority. Let $D_i(t, t+\tau)$ be the total demand of class i in the interval $(t, t+\tau]$, and $D(t, t+\tau) = D_1(t, t+\tau) + D_2(t, t+\tau)$ the total demand of both classes in the interval $(t, t+\tau]$. We denote by $F_{D_i(\tau)}(x)$ the cumulative distribution function of the total demand of class i in $[0, \tau]$ and $F_{D(\tau)}(x)$ the cumulative distribution function of the total demand of both classes in $[0, \tau]$.

In this chapter we consider fast-moving items for which is more representative and efficient to model the demand over a time period by a continuous distribution. Following Zheng (1992) we assume that the total demand of each class are represented by a nondecreasing stochastic process with continuous sample paths, and stationary and independent increments. For simplicity, we will refer to this as strictly increasing non-negative demand. This is a common assumption in stochastic inventory models (Axsäter (2007)) and is implicitly assumed in most (elementary) textbooks on inventory management. However, the assumptions of independence and continuity are conflicting; therefore, rigorously speaking, the assumption is approximate (Browne and Zipkin (1991)). Note that under stationary and independent increments, $D_i(\tau) := D_i(0, \tau) = D_i(t, t + \tau)$ for any $t \geq 0$, $i = 1, 2$.

Inventory is replenished according to a continuous review (Q, r, C) policy that operates as follows. When the inventory position falls below a reorder level r, a replenishment order for Q units is placed and arrives a fixed $L > 0$ time units later. Demand from both classes are filled as long as the on-hand inventory level is greater than the critical level C, otherwise only high priority demand is satisfied from inventory on-hand and low priority demand is backordered. If on-hand inventory level reaches zero both demands are backordered. To clear backlogged orders, we consider the threshold clearing mechanism of Deshpande et al. (2003).

Given the inventory control strategy, our objective is to find the parameters of the critical level policy that minimize the sum of the ordering cost, the holding cost and shortage costs. Let $AC(Q, r, C)$ be the average cost per unit time:

$$AC(Q, r, C) = S\frac{\mu}{Q} + h\mathbb{E}(OH(\infty)) + b_1\mathbb{E}(BO_1(\infty)) + b_2\mathbb{E}(BO_2(\infty)), \tag{4.1}$$

where μ is the average demand per unit of time; b_i is the shortage cost per unit time of class i with $b_1 > b_2 > 0$; h is the holding cost per unit and per unit time; S is the ordering cost; $\mathbb{E}(OH(\infty))$ is the steady-state on-hand inventory; and $\mathbb{E}(B_i(\infty))$ is the class i steady-state backorder, $i = 1, 2$.

In a (Q, r, C) policy with full-backorders and deterministic lead time, the inventory level is the on-hand inventory net of all backorders, i.e., $IL(t + L) = OH(t + L) - BO_1(t + L) - BO_2(t + L)$, where $IL(t + L)$ denotes the inventory level, $OH(t + L)$ denotes on-hand inventory and $BO_i(t + L)$ denotes class i backorders, $i = 1, 2$, all at time $t + L$. Furthermore, for a (Q, r, C) policy with full-backorders and deterministic lead time it is still valid that $IL(t + L) = IP(t) - D(L)$, where $IP(t)$ denotes the inventory position at time t (Deshpande et al. (2003)). Under strictly increasing non-negative demand, $IP(t)$ will be uniformly distributed on $(r, r + Q]$ in steady state and independent of lead time demand (Zheng (1992) refers to Serfozo and Stidham (1978) and Browne and Zipkin

(1991) for a detailed discussion of this assumption). Then, the on-hand inventory at time $t + L$ is $OH(t + L) = IP(t) - D(L) + BO_1(t + L) + BO_2(t + L)$. Taking the expected value and considering the limit $t \to \infty$, the expected on-hand steady-state inventory is:

$$\mathbb{E}(OH(\infty)) = \frac{Q}{2} + r - \mu L + \mathbb{E}(BO_1(\infty)) + \mathbb{E}(BO_2(\infty)). \tag{4.2}$$

4.1.1 Steady state backorders

In this section we develop expressions for backorders of the low and high priority classes in the steady state by using *state-dependent demand* approach, the inventory position and the threshold clearing mechanism. We first describe how the inventory system behaves under rationing, and the threshold clearing mechanism of Deshpande et al. (2003).

Consider an arbitrary time $t + L$. By definition, there is rationing at time $t + L$ when $C > OH(t + L) \geq IL(t + L) = IP(t) - D(t, t + L)$. Under rationing conditions at $t + L$, let t_c be the first time after t when $IP(t) - C$ demand is observed. The threshold clearing mechanism of Deshpande et al. (2003) only comes into play when backorders exist on arrival of a replenishment order and uses t_c to separate which backorders need to be cleared once the replenishment order arrives. The general rules to clear the backorders when the replenishment order arrives are:

1. If the entering replenishment batch is large enough to clear all the backorders and leave the on-hand inventory level above C, then clear all backorders,

2. Otherwise:

   2.1 Clear all backlogged demand that arrived before t_c in the order of arrival (FCFS),

   2.2 Clear any remaining backlogged class 1 demands using FCFS until either all class 1 backorders are filled, or no on-hand inventory remains,

   2.3 Carry over (i.e. continue backlog) all class 2 demands that arrive after t_c.

Note that rule 1 ensures that $OH(t) = IL(t)$ when $OH(t) \geq C$. Rule 2.2 and 2.3 mean that all remaining backorders that cannot be fulfilled by the entering replenishment batch, are carried over to be satisfied in the following replenishment arrivals.

The backorders of both low- and high-priority at time $t + L$, deducted by Deshpande et al.

(2003), are respectively:

$$BO_2(t+L) = \begin{cases} D_2(t_c, t+L) & \text{if } IP(t) - D(t, t+L) < C \\ 0 & \sim \end{cases} \tag{4.3}$$

$$BO_1(t+L) = \begin{cases} [D_1(t_c, t+L) - C]^+ \\ 0 & \sim \end{cases} \tag{4.4}$$

The inventory position $IP(t)$ does not provide enough information to determine the backorders in the steady state (Deshpande et al. (2003)). To address this lack of information, in chapter 3 we propose to use the *hitting time* approach to characterize the inventory system and we obtain exact expressions of backorders of both low- and high- priority class at steady state under the threshold clearing mechanism. The problem is that the backorders are non convex. In this chapter, we propose use an approximation to address with the lack of information and could define approximated and convex expressions of the backorders at steady state. Our approximation is intuitive, based on demand class i during $[t_c, t+L]$ is proportional to the total demand for both classes during this period, i.e.,

$$D_i(t_c, t+L) = k_i(D(t, t+L) - IP(t) + C), \quad \forall\, i = 1, 2, \tag{4.5}$$

where $D(t_c, t+L) = D(t, t+L) - IP(t) + C$, and k_i is the proportionality factor for class i. Note that k_i is is a random variable and for any proportionality factor, it must be satisfied that $k_1 + k_2 = 1$. Our approach is to consider that the proportionality factor k_i is constant. Then, replacing equation (4.5) in equations (4.3) and (4.4), and taking the expectation and conditioning on the inventory position the approximated expected backorders in the steady state of class 1 and 2 for a strictly increasing non-negative demand are, respectively:

$$\mathbb{E}(BO_2(\infty)) = \frac{k_2}{Q}\left(\beta(r - C) - \beta(r + Q - C)\right), \tag{4.6}$$

$$\mathbb{E}(BO_1(\infty)) = \frac{k_1}{Q}\left(\beta(r + C^{'}) - \beta(r + Q + C^{'})\right), \tag{4.7}$$

where: $\beta(v) = \int_v^{\infty} (x - v)(1 - F_{D(L)}(x))dx$ and $C^{'} = C\left(\frac{1}{k_1} - 1\right)$.

4.1.2 Cost optimization problem for strictly increasing non-negative demand

Replacing equation (4.2), (4.6) and (4.7) in equation (4.1) we can express the approximate average cost per unit time as:

$$\begin{aligned} AC(Q,r,C) =& S\frac{\mu}{Q} + h\left(\frac{Q}{2} + r - \mu L\right) + (b_1 + h)\frac{k_1}{Q}\left(\beta(r + C^{'}) - \beta(r + Q + C^{'})\right) \\ &+ (b_2 + h)\frac{k_2}{Q}\left(\beta(r + C^{'}) - \beta(r + Q + C^{'})\right) \end{aligned} \tag{4.8}$$

Our objective is to determine the optimal parameters of the (Q, r, C) policy that minimizes the total cost. Then, our problem for a strictly increasing non-negative demand can be written as a nonlinear optimization problem, denoted (**P0**), as follows.

Problem (**P0**):

$$\underset{Q,r,C}{\text{Min}} \quad AC(Q,r,C) \tag{4.9}$$

$$s.t \quad r \geq C \geq 0. \tag{4.10}$$

Constraint (4.10) ensures that the replenishment order is placed before the lower priority class is no longer served.

4.1.3 P0 using normal distribution as approximation of non-negative demand

A common practice in stochastic inventory models is to use the normal distribution as an approximation of non-negative demand, i.e., stochastic inventory models are formulated based on the characteristics of non-negative demand and are then implemented using a normal distribution. The problem with the normal distribution is that there is always a small probability for negative demand. Therefore, an exact result for a strictly increasing non-negative demand is only approximately true for normal demand. But, on the other hand, a normal distribution is easy to handle and the probability of negative demand can be handled through the relationship between the mean and the standard deviation. In this sense, Peterson and Silver (1979) show that normal distribution is a good approximation of non negative demand when the coefficient of variation is less than or equal to 0.5, i.e., $CV \leq 0.5$.

To solve (**P0**) using normal distribution as approximation of the non-negative demand, the expressions of backorders are required. For this, consider that each class i has identical and independent normally distributed demand per unit time, with mean $\mu_i > 0$ and variance $\sigma_i^2 > 0$, $D_i(\tau) \sim N(\mu_i \tau, \sigma_i^2 \tau)$, and $D(\tau) \sim N(\mu\tau, \sigma^2\tau)$, where $\mu = \mu_1 + \mu_2$ and $\sigma^2 = \sigma_1^2 + \sigma_2^2$. The approximated expected backorders in the steady state of class 1 and 2 using a normally distributed demand are, respectively:

$$\mathbb{E}(BO_2(\infty)) = \frac{\sigma'^2}{Q} k_2 \left[H\left(\frac{r - C - \mu'}{\sigma'}\right) - H\left(\frac{r + Q - C - \mu'}{\sigma'}\right) \right], \tag{4.11}$$

$$\mathbb{E}(BO_1(\infty)) = \frac{\sigma'^2}{Q} k_1 \left[H\left(\frac{r + C' - \mu'}{\sigma'}\right) - H\left(\frac{r + Q + C' - \mu'}{\sigma'}\right) \right], \tag{4.12}$$

where: $\mu' = \mu L$, $\sigma' = \sigma\sqrt{L}$,

$$H(x) = \int_x^{\infty} G(v)dv = \frac{1}{2}\left[(x^2+1)(1-\Phi(x)) - x\varphi(x)\right],$$

$$G(x) = \int_x^{\infty} (v - x)\varphi(v)dv = \varphi(x) - x(1 - \Phi(x)),$$

is the so-called loss function, $\Phi(x)$ is the distribution function of the standard normal distribution and $\varphi(x)$ is the density function.

As $H(x)$ is decreasing and convex Axsäter (2007)), it is easy to note that $\mathbb{E}(BO_1(\infty))$ is decreasing in r and C, while $\mathbb{E}(BO_2(\infty))$ is increasing in C and decreasing in r. The same behavior is described by Deshpande et al. (2003) for Poisson demand.

Then, the problem (**P0** + **N**) can be written as follows.

Problem (**P0+N**)

$$\begin{aligned}\min_{Q,r,C} \quad & S\frac{\mu}{Q} + h\left(\frac{Q}{2} + r - \mu L\right) + (b_1 + h)\frac{\sigma'^2}{Q} k_1 \left[H\left(\frac{r + C' - \mu'}{\sigma'}\right) - H\left(\frac{r + Q + C' - \mu'}{\sigma'}\right) \right] \\ & + (b_2 + h)\frac{\sigma'^2}{Q} k_2 \left[H\left(\frac{r - C - \mu'}{\sigma'}\right) - H\left(\frac{r + Q - C - \mu'}{\sigma'}\right) \right] \end{aligned} \tag{4.13}$$

$$s.t \quad r \geq C \geq 0. \tag{4.10}$$

4.2 Convexity of the approximate average cost per unit time $AC(Q, r, C)$

Consider the objective function of (**P0**). Since it is a nonlinear function, finding the optimal parameters of the (Q, r, C) policy is difficult, unless the objective function is convex. Clearly the first and second term of (4.8) are convex, hence the convexity of $AC(Q, r, C)$ will depend on whether the backorders are convex or not.

Proposition 7. *Backorders of class 1 and class 2 defined by equations (4.6) and (4.7) are strictly convex in Q, r and C.*

Proof. Consider the backorder in the steady-state for the continuous review (Q, r) policy:

$$\mathbb{E}(BO(\infty)) = \frac{1}{Q}\left(\beta(r) - \beta(r+Q)\right). \tag{4.14}$$

Zipkin (1986) has already proved that (4.14) is jointly convex in Q and r when $f_{D(L)} > 0$ for any $t > 0$.

The proposed approximation allows us to describe the steady-state backorders in terms of Q and a linear combination of r and C. Note that class 1 backorders depend on $r + C'$, and class 2 backorders depend on $r - C$. It is a fact that the composition of a convex function with an affine mapping preserves convexity (see Boyd and Vandenberghe (2004)). Thus, $AC(Q, r, C)$ given by (4.8) is jointly convex in Q, r and C and (**P0**) is a nonlinear convex problem. □

The above proposition applies also to the case where $F_{D(L)}(x)$ is normal. Therefore, $AC(Q, r, C)$ given by (4.13) is jointly convex in Q, r and C and (**P0 + N**) is a nonlinear convex problem.

4.3 Solution approach

For any convex optimization problem with differentiable objective and constraint functions, any points that satisfy KKT conditions are primal and dually optimal, and have zero duality gap (Boyd and Vandenberghe (2004)). Since (**P0**) and (**P0 + N**) are strictly convex optimization problems, our approach to finding a solution will be based on solving the KKT conditions.

Let us consider the optimization problem (**P0**) and (**P0** + **N**). A simple analysis of the objective function allows us to relax $Q, C \geq 0$ from (**P0**) because (i) when $Q \to 0^+$, then $AC(Q, r, C) \to \infty$, so it is not possible that $Q \geq 0$ is active in the optimum; and (ii) $\mathbb{E}(BO_1(\infty))$ decreases in C whereas $\mathbb{E}(BO_2(\infty))$ increases in C, therefore non-positive values of C would be obtained only if $b_2 \geq b_1$ which is a contradiction according to the model framework. The above relaxation applies also to (**P0** + **N**).

Regarding the reorder point, although a negative value of r does not have any practical sense, we cannot disregard the non-negativity constraint of this variable, mostly because there are some situations when this constraint is active, e.g., if the ordering cost S is too high, the optimal lot size will be so large that the optimal reorder point will have to be zero. Thus, we will solve (**P0**) and (**P0** + **N**) by only considering the $r \geq C$ and $r \geq 0$ constraints, denoted as (**P1**+) and (**P1** + **N**), respectively.

The Lagrangian function of (**P1**+) is $\mathcal{L}(Q, r, C, \lambda) = AC(Q, r, C) + \lambda_1(C - r) - \lambda_2 r$, where λ_i is a Lagrangian multiplier. Then, the KKT conditions are defined as follows:

$$
\begin{aligned}
\frac{\partial}{\partial r}\mathcal{L}(Q, r, C, \lambda_i) &= \frac{\partial}{\partial r}AC(Q, r, C) - \lambda_1 + \lambda_2 &&= 0\\
\frac{\partial}{\partial C}\mathcal{L}(Q, r, C, \lambda_i) &= \frac{\partial}{\partial C}AC(Q, r, C) + \lambda_1 &&= 0\\
\frac{\partial}{\partial Q}\mathcal{L}(Q, r, C, \lambda_i) &= \frac{\partial}{\partial Q}AC(Q, r, C) &&= 0\\
&& \lambda_1(C - r) &= 0\\
&& \lambda_2(-r) &= 0\\
&& \lambda_i &\geq 0
\end{aligned}
$$

Note that the domain of (**P1**+) is defined by linear constraints which make the KKT conditions easier to solve compared to a problem that is subjected to nonlinear constraints (e.g., service level type 1, or the fill-rate). Since these constraints are few and easy to deal with, it is straightforward to define an algorithm in terms of the activation/deactivation of them.

Let (Q^*, r^*, C^*) be the optimal solution of (**P1**+) and let (Q^u, r^u, C^u) be the solution of (**P1**+), which is unrestricted, i.e., when $\lambda_1 = \lambda_2 = 0$. The solution set is obtained from the following algorithm:

Algorithm 1 Active-constraints algorithm

Solve the KKT conditions of (**P1**+) (unrestricted)
if $r^u > C^u$ **then**
  The optimal solution is (Q^u, r^u, C^u)
else if $r^u > 0$ **then**
  (Q^*, r^*, C^*) is obtained from the KKT conditions by considering $r^* = C^*$ ($\lambda_1 > 0$ and $\lambda_2 = 0$)
else
  (Q^*, r^*, C^*) is obtained from the KKT conditions by considering $r^* = C^* = 0$ ($\lambda_i > 0$)
end if

Algorithm 1 ensures that the optimal solution will be found, however, the complexity of the equation systems derived from the KKT conditions will depend on how the demand process is modeled.

Let us now consider the (**P1 + N**) problem. Before applying algorithm 1 to find the optimal solution, it is convenient to define:

$$f_i(Q, r, C) = 1 - \frac{\sigma'}{Q}\left[G\left(\frac{a_i - \mu'}{\sigma'}\right) - G\left(\frac{a_i + Q - \mu'}{\sigma'}\right)\right], \tag{4.15}$$

where a_i is a linear combination of r and C, with $r + C'$ and $r - C$ for class 1 and class 2 demands, respectively. We can express the partial derivatives of $AC(Q, r, C)$ with respect to r and C in terms of f_i :

$$\frac{\partial}{\partial r}AC(Q, r, C) = h + k_1(b_1 + h)(f_1 - 1) + k_2(b_2 + h)(f_2 - 1) \tag{4.16}$$

$$\frac{\partial}{\partial C}AC(Q, r, C) = k_2(b_1 + h)(f_1 - 1) + k_2(b_2 + h)(1 - f_2) \quad , \tag{4.17}$$

and the partial derivative of $AC(Q, r, C)$ with respect to Q is given by the following equation:

$$\begin{aligned}\frac{\partial}{\partial Q}AC(Q, r, C) =& \frac{h}{2} - S\,\frac{\mu}{Q^2} \\ &-(b_1 + h)\frac{\sigma'^2}{Q^2}k_1\left[H\left(\frac{r + C' - \mu'}{\sigma'}\right) - H\left(\frac{r + C' + Q - \mu'}{\sigma'}\right) - \frac{Q}{\sigma'}G\left(\frac{r + C' + Q - \mu'}{\sigma'}\right)\right] \\ &-(b_2 + h)\frac{\sigma'^2}{Q^2}k_2\left[H\left(\frac{r - C - \mu'}{\sigma'}\right) - H\left(\frac{r - C + Q - \mu'}{\sigma'}\right) - \frac{Q}{\sigma'}G\left(\frac{r - C + Q - \mu'}{\sigma'}\right)\right]\end{aligned} \tag{4.18}$$

Let (Q^u, r^u, C^u) be the optimal solution of (**P1 + N**), which is unrestricted. This solution set can be obtained from the following system of equations:

$$f_1(Q, r, C) = \frac{b_1}{b_1 + h} \tag{4.19}$$

$$f_2(Q, r, C) = \frac{b_2}{b_2 + h} \tag{4.20}$$

$$\frac{\partial}{\partial Q} AC(Q, r, C) = 0 \tag{4.21}$$

Then (Q^u, r^u, C^u) is the optimal solution of $(\mathbf{P1} + \mathbf{N})$ only if it belongs to the domain of the problem, otherwise $AC(Q^u, r^u, C^u)$ is a lower bound and it is necessary to solve another system of equations to determine the optimal solution. Then, if r^u is non-negative but smaller than C, the optimal solution of $(\mathbf{P1} + \mathbf{N})$ is: $r^* = C^*$, and C^*, Q^* are obtained from equations (4.19) and (4.21), respectively. Otherwise, the optimal solution of $(\mathbf{P1} + \mathbf{N})$ is $r^* = C^* = 0$, and Q^* is obtained from equation (4.21). The above procedure can be carried out via the following algorithm.

Algorithm 2 Iterative technique to solve $(\mathbf{P1} + \mathbf{N})$

$Q^{(0)} = \sqrt{\frac{2\mu S}{h}}$
Obtain$(r + C')^{(k)}$ from 4.19
Obtain $(r - C)^{(k)}$ from 4.20
Obtain $r^{(k)}$ and $C^{(k)}$
if $r^{(k)} > C^{(k)} > 0$ **then**
    Given $(r^{(k)}, C^{(k)})$ obtain $Q^{(k+1)}$ from 4.21
else if $0 < r^{(k)} \leqslant C^{(k)}$ **then**
    Recalculate $(r^{(k)}, C^{(k)})$ from $r^{(k)} = C^{(k)}$ and 4.19
    Given $(r^{(k)}, C^{(k)})$ obtain $Q^{(k+1)}$ from 4.21
else
    Redefine $r^{(k)} = C^{(k)} = 0$
    Given $(r^{(k)}, C^{(k)})$ obtain $Q^{(k+1)}$ from 4.21
end if
if $AC(Q^{(k)}, r^{(k-1)}, C^{(k-1)}) - AC(Q^{(k+1)}, r^{(k)}, C^{(k)}) \leq \varepsilon$ **then**
    Stop
else
    go to 2
end if

Note that, $(\mathbf{P0})$ and $(\mathbf{P0} + \mathbf{N})$ can also be solved using a nonlinear convex equation.

4.4 Computational study

In this section, we present our numerical study and its results. The main objective of the computational study is to show how good is the performance of our approximation through compared the critical level policy with the separate stock and round-up policies.

The quality of the convex approximation depends on the factor of proportionality k_i, which allows us to express the demand of class i as a portion of the total demand for both classes. We propose to use the ratio of the average demand per unit time as a factor of proportionality, i.e.,

$$k_i = \frac{\mu_i}{\mu}. \tag{4.22}$$

For simplicity, we use normally distributed demand as an approximation to the non-negative demand, solving (**P0**) through the algorithm 2. Let (Q^*, r^*, C^*) be the optimal critical level policy controls of (**P0**). Once we obtain the optimal critical level policy controls we evaluate the exact expected backorders of both classes through the formulation provided by equations (3.20) and (3.21). Then, let $AC_e(Q^*, r^*, C^*)$ be average cost per unit time resulting to evaluate the exact backorders.

To evaluate the performance of our approximation, we carried out several test problems and computed the benefit of the critical level policy obtained with the proposed approach against the round-up and separate stock policies. Let AC_u and AC_s be the average cost per unit time induced by the round-up and separate stock policies respectively and $100\times(AC_u - AC_e)/AC_e$ and $100\times(AC_s - AC_e)/AC_e$ the benefit of the critical level policy against the round-up and separate stock policies respectively

In order to cover a wide range of data, we design a set of 10 experiments to evaluate the performance of our approximation and to compare the critical level policy with the separate stock and round-up policies. In each experiment we fix the shortage costs per unit time b_1 and b_2, and consider a base case with the following parameters: normal demand distributions with mean $\mu_1 = \mu_2 = 25$ and coefficient of variation $CV_1 = CV_2 = 0.2$ ($\sigma_1^2 = \sigma_2^2 = 25$), lead time $L = 5$, ordering cost $S = 300$ and holding cost per unit and unit time $h = 0.75$. We conduct experiments studying the sensitivity of the solutions to changing parameters CV_i, μ_i, S, and h. This gives a total of 135 experiments for each setting of the shortage costs.

The equation systems of Algorithm 2 was programmed by a C code using Brent-Dekker method. All test were carried on a PC with Intel Core i7 2.3 GHz processor and 16 GB RAM. The time to compute the parameters of the critical level policy are on average $8.8E-05$ seconds and in the worst case $8.8E-4$ seconds.

Our numerical results show that our approximation is able to provide good-quality solutions because the benefit of the critical level policy obtained with the proposed approach against the separate stock and round-up policies is on average 7.5% and 26.3% respectively. Table 4.1 shows the average and maximum relative benefit of the critical level policy with respect to the round-up and separate stock for the 10 settings of shortage costs and different values of S.

		Benefit (%) vs Round-up					
		$S=100$		$S=300$		$S=500$	
b_1	b_2	Average	Max	Average	Max	Average	Max
30	5	11.79	34.66	11.08	29.41	8.38	23.42
30	10	7.85	23.41	7.97	29.55	5.44	14.77
30	15	5.69	17.98	6.29	29.61	3.89	10.18
30	20	4.22	14.36	5.16	29.63	2.85	7.10
30	25	3.11	11.63	4.32	29.65	2.07	4.78
10	5	7.35	24.03	7.55	29.64	5.06	13.46
15	5	9.22	27.91	9.03	29.55	6.45	17.55
20	5	10.37	30.84	9.95	29.49	7.31	20.20
25	5	11.18	33.03	10.59	29.45	7.92	22.04
35	5	12.27	35.94	11.47	29.38	8.75	24.52

		Benefit (%) vs Separate stock					
		$S=100$		$S=300$		$S=500$	
b_1	b_2	Average	Max	Average	Max	Average	Max
30	5	26.01	31.37	26.27	30.93	26.45	30.71
30	10	25.47	31.55	26.25	30.97	26.44	30.70
30	15	25.15	31.70	26.23	31.11	26.41	30.83
30	20	24.93	31.79	26.21	31.19	26.39	30.91
30	25	24.76	31.85	26.19	31.24	26.37	30.96
10	5	26.03	32.41	26.88	31.56	26.99	31.19
15	5	26.04	31.92	26.64	31.19	26.78	30.87
20	5	26.03	31.60	26.49	30.99	26.64	30.78
25	5	26.02	31.37	26.37	30.96	26.54	30.74
35	5	25.99	31.51	26.19	30.90	26.39	30.68

Table 4.1: Benefit of the critical level vs. Round-up and Separate stock policies

Table 4.1 shows that in all experiments, the average relative benefit is greater with respect to the separate stock policy, but the maximum relative benefit is reached when comparing against the round-up policy. We also note that the relative benefit to the round-up is more sensitive and, by contrast, using two separate lot sizes and two separate reorder points causes a more homogeneous benefit. The maximum relative benefit, with respect to round-up, occurs when there is maximum difference between the shortage costs and the ordering cost is minimal ($S = 100$). Unlike, the maximum relative benefit with respect to separate stock occurs when there is minimal difference between the shortage costs and the ordering cost is minimal. As an example, table 4.2 shows the relative benefit regarding round-up and separate stock for the 135 problems of the experiment: $b_1 = 30$ and $b_2 = 5$.

			Benefit(%) vs Round-up								
			$\mu_1 = 100, \mu_2 = 25$			$\mu_1 = \mu_2 = 25$			$\mu_1 = 25, \mu_2 = 100$		
CV_1	CV_2	h	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$
0.2	0.2	0.25	2.61	1.82	1.53	3.78	29.36	2.24	8.80	6.27	5.32
		0.75	4.43	3.29	2.84	6.96	29.40	4.52	15.12	11.49	10.02
		1.25	5.59	4.29	3.75	9.14	29.41	6.25	19.13	15.06	13.34
0.4	0.4	0.25	4.47	3.31	2.82	6.13	4.35	3.67	13.20	9.99	8.60
		0.75	6.89	5.50	4.83	10.37	7.85	6.81	20.57	16.80	14.94
		1.25	8.32	6.88	6.13	13.04	10.21	8.99	24.89	21.05	19.04
0.6	0.6	0.25	5.96	4.66	4.05	8.03	5.89	5.01	16.10	12.84	11.27
		0.75	8.69	7.32	6.58	12.83	10.11	8.86	23.74	20.41	18.56
		1.25	10.28	8.93	8.16	15.72	12.83	11.42	28.15	24.91	23.02
0.6	0.2	0.25	5.77	4.49	3.89	6.16	4.39	3.71	9.40	6.77	5.75
		0.75	8.43	7.08	6.35	10.27	7.83	6.79	15.75	12.15	10.61
		1.25	9.99	8.65	7.88	12.84	10.14	8.93	19.70	15.75	13.98
0.2	0.6	0.25	3.26	2.30	1.94	7.27	5.20	4.40	16.39	13.02	11.41
		0.75	5.37	4.06	3.50	12.08	9.25	8.05	24.24	20.77	18.86
		1.25	6.68	5.22	4.57	15.04	11.93	10.54	28.76	25.38	23.42

			Benefit(%) vs Separate stock								
			$\mu_1 = 100, \mu_2 = 25$			$\mu_1 = \mu_2 = 25$			$\mu_1 = 25, \mu_2 = 100$		
CV_1	CV_2	h	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$	$S = 100$	$S = 300$	$S = 500$
0.2	0.2	0.25	24.18	24.73	24.90	29.42	29.36	29.34	24.84	25.27	25.39
		0.75	23.61	24.31	24.53	29.50	29.40	29.36	24.88	25.49	25.66
		1.25	23.28	24.03	24.26	29.55	29.41	29.36	24.99	25.72	25.92
0.4	0.4	0.25	23.62	24.36	24.61	29.78	29.59	29.52	24.38	25.04	25.24
		0.75	23.14	23.98	24.28	30.09	29.79	29.68	24.44	25.28	25.55
		1.25	22.98	23.80	24.11	30.30	29.93	29.78	24.58	25.55	25.86
0.6	0.6	0.25	23.48	24.23	24.52	30.25	29.92	29.79	24.29	25.00	25.24
		0.75	23.27	24.02	24.33	30.80	30.33	30.14	24.55	25.40	25.70
		1.25	23.34	24.01	24.30	31.18	30.62	30.38	24.90	25.79	26.12
0.6	0.2	0.25	21.28	23.02	23.65	29.79	29.81	29.77	28.35	27.35	26.98
		0.75	20.79	22.67	23.39	30.59	30.46	30.33	30.18	28.70	28.12
		1.25	20.78	22.63	23.36	31.18	30.93	30.71	31.37	29.63	28.91
0.2	0.6	0.25	25.18	25.12	25.11	27.25	28.14	28.44	16.39	22.49	23.32
		0.75	24.63	24.59	24.61	26.23	27.57	28.02	24.24	21.82	22.95
		1.25	24.22	24.20	24.23	25.59	27.22	27.78	28.76	21.58	22.89

Table 4.2: Benefit(%) vs. Round-up and Separate stock when $b_1 = 30$ and $b_2 = 5$

The pattern of the maximum relative benefit regarding round-up policy, observed in table 4.2, is repeated for all ten experiments, i.e., the maximum benefit occurs when the class 2 dominates on mean and variance ($\mu_2 = 100$, $CV_2 = 0.6$), the ordering cost is minimal ($S = 100$) and the holding cost per unit and unit time is maximum ($h = 1.25$). Clearly, the round-up policy is highly inefficient when the class 2 dominates mean and variance, because under this situation, this policy provides too much inventory to the low priority class causing a high reorder point and therefore a high cost. On the other hand, when ordering cost is low and holding cost per unit and unit time is high, batch sizes are small and the expected backorder increases. We observe that the expected backorders induced by the critical level are greater than those induced by the round-up policy, but its effect on cost is relatively low compared with the effect of the reorder point. Note that, as Deshpande et al. (2003) observed, the relative benefit regarding Round-up is decreasing in S.

4.5 Conclusions

In this chapter we analyzed the constant critical level policy for fast-moving items when the inventory system faced random demands from two customer classes (high and low priority). The inventory system operated under a continuous review (Q, r) policy, with a critical threshold value C, full-backorders and a deterministic lead time. Penalty cost of backorders of the high-priority class were greater than the low-priority class, and the demand of each class was characterized by a strictly increasing non negative demand. We also characterized the demand of each class with a normal distribution, which acted as an approximation of non-negative demand. Using the *state-dependent demand* approach and the *threshold clearing mechanism*, we obtained an approximation for on-hand inventory based on a convex approximation of backorders. The approximation considered that the demand class was proportional to the total demand for both classes. We then proposed a nonlinear cost optimization problem with convex objective functions to determine the parameters of the critical level policy. Given the convexity of the cost optimization problem we proposed a system of equations to solve it. Once we obtained the optimal parameters of the critical level policy, we evaluate the exact expected backorders which allows us to obtain the exact average cost per unit time.

Our numerical results show that our approximation is able to provide good-quality solutions because the benefit of the critical level policy obtained with the proposed approach against the separate stock and round-up policies is on average 7.5% and 26.3% respectively.

There are a number of questions and issues left for future research. The first is to find a better proportionality factor, while the second is to expand our results to more than two classes. Third is the joint optimization of the order quantity, reorder point and critical level with an exact formulation of backorders. Finally, another line of future work is to investigate the equivalence between shortage costs and fill-rate service levels.

5 | Joint location-inventory problem with two demand classes

This chapter analyzes the design of a distribution network for fast-moving items able to provide differentiated service levels in terms of product availability for two demand classes (high and low priority) using a critical level policy. The model is formulated as a MINLP with chance constraints for which we propose a heuristic to solve it. Although the heuristic does not guarantee an optimal solution, our computational experiments have shown that it provides good-quality solutions that are on average 0.8% and at worst 2.7% from the optimal solution.

5.1 Model formulation

Our location-inventory model with differentiated service levels can be stated as follows. Consider the design of a distribution network consisting of an external supplier and a set of J candidate sites for locating DCs which must supply a set I of retailers. These retailers could be customers or markets, but for convenience we denote them as retailers in the rest of this chapter. We assume that the location of the external supplier, site candidates and retailers locations are known and that the supplier and DCs are uncapacitated. In this distribution network there are two categories of retailers or demand classes (high and low priority). The high priority retailers (class 1) require high service level and the low priority retailers (class 2) require lower service level. A retailer can be assigned to a single demand class, and we define $N_k = \{i \in I \mid i \text{ is class } k\}$, with $k = 1, 2$, as the set of retailers of class k. We also assume that the class of each retailer is known. The demands per unit time at each retailer are independent and normally distributed with mean $\mu_i > 0$ and variance $\sigma_i^2 > 0$. The problem is to determine the optimal number of DCs, their locations, the retailers assigned to each DC, how much inventory to keep at each of them and how to meet the preset service level for each demand class so as to minimize the total system cost.

To provide differentiated service levels we assume at each DC a continuous review (Q, r, C) policy with full-backorder and deterministic lead time, operating as follows. When the inventory position at DC j falls below a reorder level r_j, a replenishment order for Q_j units is placed and arrives $L_j > 0$ time units later. Demand from both classes is satisfied as long as the inventory level is greater than the critical level C_j, otherwise only high priority demand is satisfied from on-hand inventory and low priority demand is backordered. If on-hand inventory level reaches zero, both demands are backordered. If a DC j provides only retailers belonging to one class demand, the critical level is zero, i.e., $C_j = 0$, and therefore, the rationing policy becomes the traditional continuous review (Q, r) policy.

As Daskin et al. (2002) and Shen et al. (2003) do, we assume the replenishment order Q_j is determined using an economic order quantity model (EOQ) and the steady-state backorders are negligible. Hence, there are four types of decision variables in our model: the reorder point in a candidate DC j, r_j; the critical level in the candidate DC j, C_j; the location variable, X_j; and the assignment variable, Y_{ij}. In particular, the last two variables are defined as:

$$X_j = \begin{cases} 1 & \text{if we locate a DC at candidate site } j \\ 0 & \sim \end{cases}$$

$$Y_{ij} = \begin{cases} 1 & \text{if demand at retailer } i \text{ is assigned to a DC at candidate site } j \\ 0 & \sim \end{cases}$$

Once we define the allocation variable Y_{ij}, we can characterize the demand at candidate DC j. Let $D_{kj}(\tau)$ be the total demand of class k during an interval of length τ at DC j and $D_j(\tau) = D_{1j}(\tau) + D_{2j}(\tau)$ be the total demand of both classes during an interval of length τ at DC j. As retailers, demand per unit time are independent and normally distributed, $D_{kj}(\tau)$ is also normally distributed with mean $\tau\mu_{kj}$ and variance $\tau\sigma_{kj}^2$, where $\mu_{kj} = \sum_{i \in N_k} \mu_i Y_{ij} \geq 0$ and $\sigma_{kj}^2 = \sum_{i \in N_k} \sigma_i^2 Y_{ij} \geq 0$. Furthermore, $D_j(\tau)$ is normally distributed with mean $\tau\mu_j$ and variance $\tau\sigma_j^2$, where $\mu_j = \mu_{1j} + \mu_{2j} = \sum_i \mu_i Y_{ij}$ and $\sigma_j^2 = \sigma_{1j}^2 + \sigma_{2j}^2 = \sum_i \sigma_i^2 Y_{ij}$. In appendix B.1, we present a glossary of terms.

5.1.1 Service level type I under rationing

In this chapter, the service level provided by a DC j to the class k is measured by the probability of satisfying the entire demand of class k assigned to him during a replenishment cycle from it on hand inventory, i.e., service level type I, which does not depend of the replenishment batch quantity Q_j.

For a single class demand, service level type I is defined as the probability of no stockout per order cycle (Axsäter (2007), page 94), i.e., the probability to satisfying the entire demand during a replenishment cycle. Mathematical formulation of these measure depends of the type of the inventory system. In a traditional (Q, r) policy, the service level type I is the probability that total demand during the lead time is less than or equal to the reorder point (Axsäter (2007), page 97), equivalent to the probability of no-stock out or satisfy the total demand during lead time. In a continuous review (Q, r, C) policy, we developed expressions for the service level type I for high and low priority class under strictly increasing non-negative demand using normally distributed demand (chapter 3), i.e., we formulate expressions for service level type I of each class demand considering non-negative demand, and then imposing the normally distributed demand as an approximation. In order to determine the operational characteristics of the inventory system we use a hitting time approach. In our work, the hitting time at DC j, $\tau_{H,D_j}^{r_j-C_j}$, is defined as the amount of time that elapses from the moment an order is placed in DC j until the time at which the critical level C_j is reached for the first time, i.e., $\tau_{H,D_j}^{r_j-C_j} = \inf\{\tau > 0 \mid D_j(\tau) > r_j - C_j\}$. The subscript H is used to remind the reader we refer to a hitting time, in this case the first time that demand D_j accumulates an amount of $r_j - C_j$.

Let $\alpha_k^j(r_j, C_j, Y_{ij})$ be the service level type I provided by the DC j to class k, and $\overline{\alpha}_k$ the preset service level for class k, where $\overline{\alpha}_1 > \overline{\alpha}_2$. Using the expressions developed in chapter 3, the service level type I provided by the DC j to the high and low priority class, under strictly increasing non-negative demand, are:

$$\alpha_1^j(r_j, C_j, Y_{ij}) = \mathbb{P}(D_j(L_j) \leq r_j - C_j) + \mathbb{P}\left(D_{1j}(L_j - \tau_{H,D_j}^{r_j-C_j}) \leq C_j \,\cap\, \tau_{H,D_j}^{r_j-C_j} < L_j\right), \tag{5.1}$$

$$\alpha_2^j(r_j, C_j, Y_{ij}) = \mathbb{P}(D_j(L_j) \leq r_j - C_j), \tag{5.2}$$

where the first term of equation (5.1) is the probability that rationing does not exist in the lead time of DC j; and second term of equation (5.1) is the probability of rationing occurs in DC j and the class 1 demand during this period not reach the critical level C_j.

In chapter 3 we assume that a single DC serves both types of demand. In our case, a DC j could serve both types of demand or only one type of demand. Therefore it is necessary to verify that the equations (5.1) and (5.2) make sense in the case that the DC j is assigned a single type of demand.

Proposition 8. *Under strictly increasing non-negative demand, equations (5.1) and (5.2) are general expressions for the service level type I provided to the high and low priority classes respectively.*

Proof. The proof is detailed in the appendix B.2. □

Using normally distributed demand and conditioning on the hitting time, the service levels provided by the candidate DC j to the high and low priority class are:

$$\alpha_1^j(r_j, C_j, Y_{ij}) = \int_0^{L_j} \Phi\left(\frac{C_j - (L_j - \tau)\sum_{i\in N_1}\mu_i Y_{ij}}{\sqrt{(L_j-\tau)\sum_{i\in N_1}\sigma_i^2 Y_{ij}}}\right) f_{H,D_j}^{r_j-C_j,Y_{ij}}(\tau)\,d\tau + \Phi\left(\frac{r_j - C_j - L_j\sum_i \mu_i Y_{ij}}{\sqrt{L_j\sum_i \sigma_i^2 Y_{ij}}}\right), \tag{5.3}$$

$$\alpha_2^j(r_j, C_j, Y_{ij}) = \Phi\left(\frac{r_j - C_j - L_j\sum_i \mu_i Y_{ij}}{\sqrt{L_j\sum_i \sigma_i^2 Y_{ij}}}\right), \tag{5.4}$$

where $\Phi(x)$ is the distribution function of the standard normal distribution,

$$f_{H,D_j}^{r_j-C_j,Y_{ij}}(\tau) = \frac{1}{\sqrt{\tau\sum_i \sigma_i^2 Y_{ij}}}\left(\frac{r_j - C_j + \tau\sum_i \mu_i Y_{ij}}{2\tau}\right)\varphi\left(\frac{r_j - C_j - \tau\sum_i \mu_i Y_{ij}}{\sqrt{\tau\sum_i \sigma_i^2 Y_{ij}}}\right), \tag{5.5}$$

is the density distribution of the hitting time $\tau_{H,D_j}^{r_j-C_j}$ using normally distributed demand and $\varphi(x)$ is the density function of the standard normal distribution. The superscripts indicates the dependence of the density distribution of the hitting time with respect to $r_j - C_j$ and assignment variable Y_{ij}, and the subscript D_j represent the total demand of both classes and H the hitting time.

5.1.2 Cost function

There is a fixed setup cost f_j of opening each distribution center. Each DC can serve more than one retailer, but each retailer should be only assigned to exactly one DC. The ordering cost from distribution center j is S_j. Linear transportation costs are incurred for shipment from the external supplier to distribution center j with unit cost a_j and from distribution center j to retailer i with unit cost d_{ij}. With this notation, the average cost per unit time at DC j is:

$$AC_j(r_j, C_j, Y_{ij}) = S_j \frac{\sum_i \mu_i Y_{ij}}{Q_j} + a_j \sum_i \mu_i Y_{ij} + \sum_i d_{ij}\mu_i Y_{ij} + h_j \left(\frac{Q_j}{2} + r_j - L_j \sum_i \mu_i Y_{ij} \right). \tag{5.6}$$

The first term of equation (5.6) is the ordering cost per unit time. The second and third term are the supply and distribution costs per unit time respectively. As we assume negligible backorders, the fourth term is approximated the holding cost per unit time. Each distribution center determines the replenishment order Q_j using an EOQ model, i.e.,

$$Q_j = \sqrt{\frac{2S_j}{h_j} \sum_i \mu_i Y_{ij}}\,. \tag{5.7}$$

Then, replacing equation (5.7) into equation (5.6), the average cost per unit time at DC j is:

$$AC_j(r_j, C_j, Y_{ij}) = \sum_i \hat{d}_{ij} Y_{ij} + k_j \sqrt{\sum_i \mu_i Y_{ij}} + h_j r_j\,, \tag{5.8}$$

where $k_j = \sqrt{2h_j S_j}$ and $\hat{d}_{ij} = (a_j + d_{ij} - h_j L_j)\mu_i$.

5.1.3 Problem formulation

We can formulate an integrated location-inventory model with differentiated service levels using normally distributed demand as an MINLP problem with constraints on service probability (non-convex) and nonlinear objective function including non-convexity in the assignments variables, denoted (**P0**), as follows.

Problem (**P0**):

$$\min_{X,Y,r,C} \quad \sum_{j\in J}\left\{f_jX_j + \sum_i \hat{d}_{ij}Y_{ij} + k_j\sqrt{\sum_i \mu_i Y_{ij}} + h_j r_j\right\} \tag{5.9}$$

$$\text{s.t:} \quad \int_0^{L_j} \Phi\left(\frac{C_j - (L_j-\tau)\sum_{i\in N_1}\mu_i Y_{ij}}{\sqrt{(L_j-\tau)\sum_{i\in N_1}\sigma_i^2 Y_{ij}}}\right) f_{H,D_j}^{r_j-C_j,Y_{ij}}(\tau)\, d\tau + \Phi\left(\frac{r_j - C_j - L_j\sum_i \mu_i Y_{ij}}{\sqrt{L_j\sum_i \sigma_i^2 Y_{ij}}}\right) \ge \overline{\alpha}_1 \quad \forall j\in J, \tag{5.10}$$

$$r_j - C_j \ge L_j\sum_i \mu_i Y_{ij} + z_{\overline{\alpha}_2}\sqrt{L_j\sum_i \sigma_i^2 Y_{ij}} \qquad \forall j\in J, \tag{5.11}$$

$$\sum_{j\in J} Y_{ij} = 1 \qquad \forall i\in I, \tag{5.12}$$

$$Y_{ij} \le X_j \qquad \forall i\in I, \forall j\in J, \tag{5.13}$$

$$r_j \ge C_j \ge 0 \qquad \forall j\in J, \tag{5.14}$$

$$X_j,\ Y_{ij} \in \{0,1\} \qquad \forall i\in I, \forall j\in J, \tag{5.15}$$

where $z_{\overline{\alpha}_2}$ is the inverse standard normal distribution for a preset service level $\overline{\alpha}_2$ and the $f_{H,D_j}^{r_j-C_j,Y_{ij}}(\tau)$ is given by equation (5.5).

The objective is to minimize the total steady state cost per unit time including location fixed cost, ordering costs, supply cost from supplier to DCs, distribution cost from DCs to retailers, and holding cost at each DC. Constraints (5.10) and (5.11) ensure, at each DC, the fulfillment of the preset service level for the high and low priority class respectively. Constraint (5.12) ensures that each retailer is assigned to exactly one DC. Constraint (5.13) stipulates that retailers can only be assigned to open DCs. Constraint (5.14) ensures, at each DC, that the replenishment order be placed before the lower priority class is no longer served. Finally, constraint (5.15) is an integrality constraint.

5.2 Solution approach

Consider the joint location-inventory problem described by Shen et al. (2003), but without relying on the assumption that each retail has identical variance-to-mean ratio. This problem assumes that the distribution network is dominated by a continuous review (Q, r) policy, deterministic lead time and full-backorders. These authors considered a single demand class, i.e., to the distribution network providing a unique service level. Let $\overline{\alpha}$ be this unique preset service level. Using the notation above the model of Shen et al. (2003) is expressed as:

Problem (**P1**):

$$\min_{X,Y} \quad \sum_{j\in J}\left\{f_jX_j+\sum_i(\hat{d}_{ij}+h_jL_j\mu_i)Y_{ij}+k_j\sqrt{\sum_i\mu_iY_{ij}}+z_{\overline{\alpha}}h_j\sqrt{L_j}\sqrt{\sum_i\sigma_i^2Y_{ij}}\right\} \tag{5.16}$$

s.t: (5.12), (5.13), (5.15).

It is easy to show that the problem (**P1**) is a relaxation of (**P0**) when $\overline{\alpha}=\overline{\alpha}_2$. Therefore, the optimal solution of the problem (**P1**) when the level of service is the lowest, is a lower bound (LB) of problem (**P0**).

Also, if we solve (**P1**), and then use the resulting configuration in terms of location (**X**-variables) and allocation (**Y**-variables), we can obtain the optimal remaining feasible variables corresponding to this configuration, for each DC, always considering the actual proportion of customers requiring each class of service ($\overline{\alpha}_2$ and $\overline{\alpha}_1$). The result is a feasible solution and, hence, an upper bound (UB) for the problem (**P0**). Thus, we now have a lower bound and an upper bound of the original problem with two classes of service, obtained by solving two problems with a single class of service, and completing the solution of the second one by including both service classes.

Note that for fixed **X**-variables and **Y**-variables that satisfy the constraint (5.12), (5.13) and (5.15), the problem (**P0**) reduces to determine the optimal parameters of the critical level policy at all installed DCs, i.e., for all $X_j=1$ we must solve the following service level problem SLP(j) using normally distributed demand:

Problem **SLP**(j):

$$\min_{r_j,C_j} \quad r_j \tag{5.17}$$

s.t: (5.10), (5.11), (5.14).

In chapter 3 we characterize under mild assumptions, the optimal solution of the problem **SLP(j)** using normally distributed demand, when a single DC satisfies the demand of both demand classes, i.e., when $\mu_{kj},\sigma_{kj}^2>0$.

Consider now that we solve the same two problems (**P1**) and **SLP(j)** for an increasing value of $\overline{\alpha}$, starting at $\overline{\alpha}_2$. Let $X^*_{\overline{\alpha}}$, $Y^*_{\overline{\alpha}}$ be the optimal location and assignments variables of the problem

(**P1**) given a service level $\overline{\alpha}$; $FO_{P1}(\overline{\alpha})$ be the objective function of problem (**P1**) given a service level $\overline{\alpha}$; and $FO_{P0}(X^*_{\overline{\alpha}}, Y^*_{\overline{\alpha}})$ be the objective function of problem (**P0**) given the optimal network configuration of the problem (**P1**) and a service level $\overline{\alpha}$. We propose a simple heuristic to find a better upper bound of the problem (**P0**) based on the following four properties:

(1) $LB = FO_{P1}(\overline{\alpha}_2)$, $UB = FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2})$ and $UB > LB$;

(2) $FO_{P1}(\overline{\alpha})$ is strictly increasing in $\overline{\alpha}$ for all $\overline{\alpha} \in [\overline{\alpha}_2, \overline{\alpha}_1]$ with $\overline{\alpha}_2 \geq 0.5$ because increasing the requirements in quality of service tightens the feasible space. However, as $\overline{\alpha}$ increases, there is no guarantee that $FO_{P1}(\overline{\alpha})$ is a lower bound anymore;

(3) Recall that the feasible solution $FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})$ is computed by finding variables X and Y as if there were only customers requiring service at level $\overline{\alpha}_1$, and completed using the actual distribution of classes of customers. Then, $FO_{P1}(\overline{\alpha}_1) \geq FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})$, because a *global round-up policy* induces a higher cost than a critical level policy. In other words, providing all customers with the highest class of service $\overline{\alpha}_1$ is costlier than having some of them with the lowest class of service $\overline{\alpha}_2$.

(4) As we increase the service level $\overline{\alpha}$ and solve the problem (**P1**) one or more of the following events may occur:

   (i) the configuration $(X^*_{\overline{\alpha}_2},\ Y^*_{\overline{\alpha}_2})$ remains constant for any $\overline{\alpha} \in [\overline{\alpha}_2, \overline{\alpha}_1]$;

   (ii) the optimal network configuration $(X^*_{\overline{\alpha}},\ Y^*_{\overline{\alpha}})$ changes due to the reallocation of demand without changing the number of open DCs; and

   (iii) the optimal network configuration $(X^*_{\overline{\alpha}},\ Y^*_{\overline{\alpha}})$ tends to make pooling, closing one or more DCs and reassigning demand.

Our improvement heuristic exploits the risk pooling effect to increase the service level $\overline{\alpha}$ in the interval $[\overline{\alpha}_2, \overline{\alpha}_1]$. If pooling happens at some $\hat{\alpha} \in [\overline{\alpha}_2, \overline{\alpha}_1]$ and the sum of the savings on holding and fixed costs are greater than the increase in transportation costs at that point, then $FO_{P0}(X^*_{\hat{\alpha}-\varepsilon}, Y^*_{\hat{\alpha}-\varepsilon}) > FO_{P0}(X^*_{\hat{\alpha}}, Y^*_{\hat{\alpha}})$. A heuristic, to find this $\hat{\alpha}$ (if exist) based on the systematic increase in the service level $\overline{\alpha}$ can be costly in terms of computational time because each increased service level $\overline{\alpha}$ means solving the problems (**P0**) and **SLP(j)**. We propose to evaluate $FO_{P0}(X^*_{\overline{\alpha}}, Y^*_{\overline{\alpha}})$ only in $\overline{\alpha} = \overline{\alpha}_1$ and $\overline{\alpha} = \overline{\alpha}_2$. Let FO^*_{P0} the objective value of the best solution found to (**P0**), then:

- if $FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2}) \leq FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})$, an increase in the service level $\overline{\alpha}$ in the interval $[\overline{\alpha}_2, \overline{\alpha}_1]$ produces no improvement in the initial solution (UB), which becomes the best solution found, i.e., $FO^*_{P0} = FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2})$;

- if $FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2}) > FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})$, there is improvement in the initial solution and the second solution becomes the best solution found, i.e., $FO^*_{P0} = FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})$;

i.e., $FO^*_{P0} = \min\{FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2}), FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})\}$.

5.2.1 Solution characterization for the problem (P1)

The square root term in the objective function of problem (**P1**) can give rise to difficulties in the optimization procedure. When the DC j is not selected, both square root terms would take a value of 0, which leads to unbounded gradients in the NLP optimization and hence numerical difficulties. Thus, we reformulate the problem (**P1**) in order to eliminate the square root terms. We first introduce two sets of nonnegative continuous variables, $Z1_j$ and $Z2_j$, to represent the square root terms in the objective function:

$$Z1_j^2 = \sum_i \mu_i Y_{ij}^2, \quad \forall j \in J \tag{5.18}$$

$$Z2_j^2 = \sum_i (\sigma_i Y_{ij})^2, \quad \forall j \in J \tag{5.19}$$

$$Z1_j, Z2_j \geq 0, \quad \forall j \in J \tag{5.20}$$

Because the nonnegative variables $Z1_j$ and $Z2_j$ are introduced in the objective function with positive coefficients, and this problem is a minimization problem, the equations (5.18) and (5.19) can be further relaxed as the following inequalities:

$$Z1_j^2 \geq \sum_i \mu_i Y_{ij}^2, \quad \forall j \in J \tag{5.21}$$

$$Z2_j^2 \geq \sum_i (\sigma_i Y_{ij})^2, \quad \forall j \in J \tag{5.22}$$

Note that constraints (5.21) and (5.22) with constraint (5.20) define second-order cone constraints. Thus, the reformulated problem (**P1**) can be expressed as the following MINLP problem with second-order cone constraints denoted as (**P2**):

Problem (**P2**):

$$\min_{X,Y,Z1,Z2} \quad \sum_{j \in J} \left\{ f_j X_j + \sum_i (\hat{d}_{ij} + h_j L_j \mu_i) Y_{ij} + k_j Z1_j + z_{\overline{\alpha}} h_j \sqrt{L_j} Z2_j \right\} \tag{5.23}$$

$$\text{s.t:} \quad (5.21), (5.22), (5.12), (5.13), (5.15), (5.20).$$

Problem (**P2**) can be trivially shown to be equivalent to problem (**P1**) but it has a linear objetive function and second-order cone constraints (5.21) and (5.22). We can solve this problem using CPLEX 12.4., which handles second-order cone constraints in an efficient way.

5.2.2 Solution characterization for SLP(j) using normally distributed demand

In this section we suppress the j subscript in order to simplify the notation. In chapter 3 we characterized under mild assumptions, the optimal solution of the problem **SLP(j)** under normally distributed demand, when a single DC satisfies demand of both classes as follows:

For, $\mu_1, \sigma_1^2, \mu_2, \sigma_2^2 > 0$ and $\overline{\alpha}_2 \in [0.5, 1)$, the optimal parameters of the critical level policy are obtained from the following system of equations:

(a) If $\alpha_1(r_2^0, 0) < \overline{\alpha}_1$:

$$r^* - C^* = \mu L + z_{\overline{\alpha}_2}\, \sigma \sqrt{L}, \tag{5.24}$$

$$\int_0^L \Phi\left(\frac{C^* - \mu_1(L-\tau)}{\sqrt{(L-\tau)\sigma_1^2}}\right) f_{H,D}^{r^*-C^*}(\tau)\, d\tau = \overline{\alpha}_1 - \overline{\alpha}_2, \tag{5.25}$$

where $r_2^0 = \mu L + z_{\overline{\alpha}_2}\, \sigma \sqrt{L}$ and the service levels provided to each class are equal to their preset levels, i.e., $\alpha_i(r^*, C^*) = \overline{\alpha}_i,\ i = 1, 2$.

(b) If $\alpha_1(r_2^0, 0) \geq \overline{\alpha}_1$:

$$C^* = 0, \tag{5.26}$$

$$r^* = \mu L + z_{\overline{\alpha}_2}\, \sigma \sqrt{L}, \tag{5.27}$$

and service levels provided to each class are: $\alpha_1(r^*, 0) \geq \overline{\alpha}_1$ and $\alpha_2(r^*, 0) = \overline{\alpha}_2$ for high and low priority class respectively.

In our network design, a candidate DC j can provide both demand classes, one or none. The following proposition indicates the optimal parameters of rationing policy when a DC provides only a single demand class.

Proposition 9. *Under normally distributed demand and single class demand, the optimal parameters of the critical level policy are:*

(a) If $\mu_1 = \sigma_1^2 = 0$ and $\mu_2, \sigma_2^2 > 0$, the optimal parameters of the critical level policy are: $C^* = 0$, $r^* = r_2^0 = \mu_2 L + z_{\overline{\alpha}_2} \sigma_2 \sqrt{L}$, and the service levels provided to the high and low priority class are $\alpha_1(r^*, C^*) = 1$ and $\alpha_2(r^*, C^*) = \overline{\alpha}_2$ respectively.

(b) If $\mu_1, \sigma_1^2 > 0$ and $\mu_2 = \sigma_2^2 = 0$, the optimal parameters of the critical level policy are $C^* = 0$ and $r^* = r_1^0$ solution of $\alpha_1(r^*, 0) = \overline{\alpha}_1$, i.e.,

$$\int_0^L \Phi\left(\frac{-\mu_1(L-\tau)}{\sqrt{(L-\tau)\sigma_1^2}}\right) f_{H,D_1}^{r^*}(\tau)\, d\tau + \Phi\left(\frac{r^* - \mu_1 L}{\sqrt{\sigma_1^2 L}}\right) = \overline{\alpha}_1, \tag{5.28}$$

and the service levels provided to the high and low priority class are $\alpha_1(r^*, C^*) = \overline{\alpha}_1$ and $\alpha_2(r^*, C^*) > \overline{\alpha}_2$ respectively.

Proof. The proof is detailed in the appendix B.3 □

Note that under no demand for one of the classes $C^* = 0$. Therefore, the (Q, r, C) policy is equivalent to the traditional (Q, r) policy.

5.3 Computational Study

In order to illustrate the applicability and evaluate the performance of our solution approach in terms of quality solution (optimality gap) and computational time, we carried out computational experiments for instances with 49 and 88 nodes from Daskin (2011). We generated several test problems for each data set in which we compare our solution approach with the Global Round-up policy. We denote these instances as *test sets*. In all cases, each retailer location is also a candidate DC location, i.e., there are as many candidate DC locations as retailer locations for each instance.

The test problems were generated around a base case with the following parameters: service level requirements $\overline{\alpha}_1 = 0.975$ and $\overline{\alpha}_2 = 0.75$; cost per unit to ship between retailer i and candidate DC site j, d_{ij} equal to the distance between retailer i and candidate DC j multiplied by a transport rate $c_{ij} = 0.01,\ \forall i, j$; and holding cost per unit and unit time at candidate DC site j, $h_j = 0.25,\ \forall j \in J$. Furthermore, all the test problems used the following common criterion and parameters: demand per unit time at each retailer is normally distributed with mean $\mu_i = U[10, 50]$ and coefficient of variation $CV_i = U[0.1, 0.4]$; class of the retail i, $n_i = \{1, 2\}$ with discrete uniform distribution; fixed (per unit time) cost of locating a DC at candidate site j, $f_j = U[200, 300]$; cost per unit to ship between external supplier and candidate DC site j, $a_j = 0.5,\ \forall j \in J$; ordering cost from candidate DC site j, $S_j = 1000, \forall j \in J$; and lead time, $L_j = \{2, 3, 4\}$ with discrete uniform distribution.

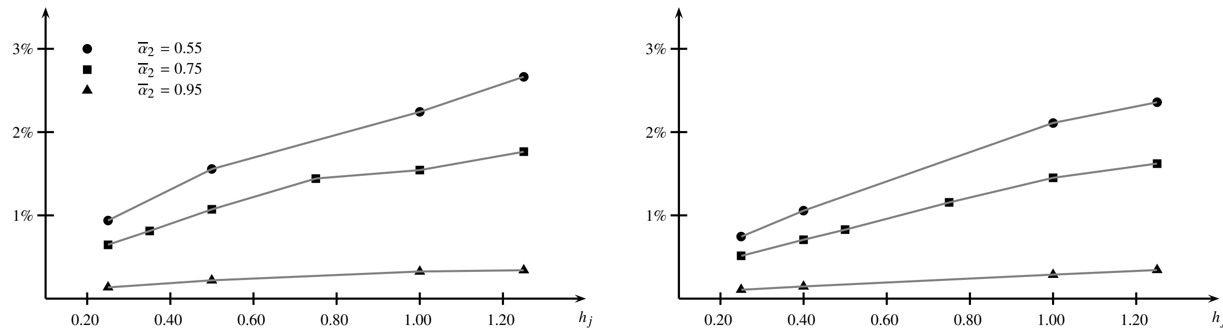
Problem (**P2**) was modeled with AMPL and solved with CPLEX 12.4. The equation systems (5.24)-(5.25) and proposition 9, solutions of problem $SLP(j)$, were programmed and solved by a C code. We integrate both codes through an AMPL script and the shell command to execute the C code. The time limit was set for 10800 seconds. All test were carried on a PC with Intel Core i7 2.3 GHz processor and 16 GB RAM.

We solved 63 problems, 30 for 49 nodes and 33 for 88 nodes. In each problem, we changed parameters relative to the base case. In particular the modified parameters were: the preset service level for low priority class ($\overline{\alpha}_2$), the holding cost per unit and unit time (h_j), and the transport rate (c_{ij}). Table B.2 (appendix B.4) shows the data-set used, the parameter modified from the base case; results of location-inventory model with differentiated service levels using critical level policy; results of the global round-up policy; and the relative difference between the total costs induced by critical level and global round-up policies.

Regarding the difficult to solve the problem we have the following comments derived from the computational experiments in table B.2 (appendix B.4):

- as expected, as the problems grow larger, it becomes more difficult to solve them;
- as the holding cost per unit and unit time and/or the preset service level increases, the problem (**P2**) becomes harder to solve. This is because, higher values of h_j and $z_{\overline{\alpha}}$ assign more weight on the nonlinear terms of the objective function of (**P2**).

We measure, for our approach, the relative optimally gap between the lower bound (LB) and the best solution found, i.e., $Gap(\%) = 100 \times (FO^*_{P0} - LB)/LB$. Figure 5.1 show for each data set, how the relative optimally gap change as the holding cost per unit and unit time (h_j) and the preset service level of the low priority class ($\overline{\alpha}_2$) change.



(a) 49-node **(b)** 88-nodes

Figure 5.1: Relative Gap with $\overline{\alpha}_1 = 0.975$, $c_{ij} = 0.01$

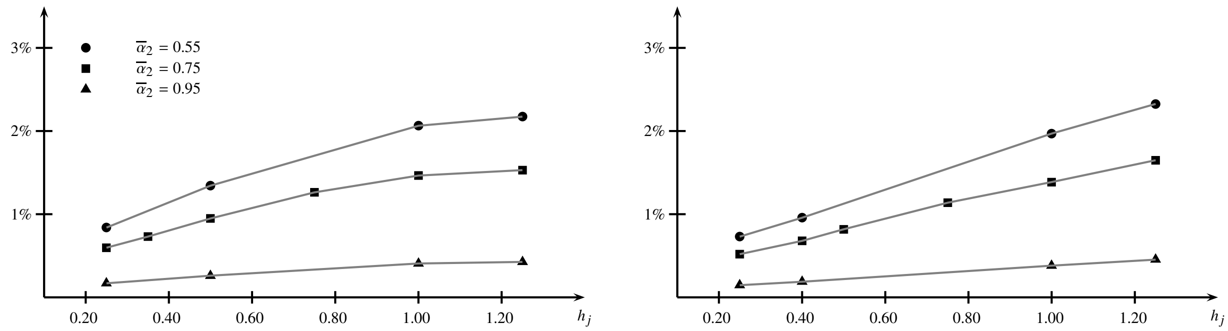
From Figure 5.1, we can see that relative optimality gap increase when the holding cost per unit and unit time and $\overline{\alpha}_1 - \overline{\alpha}_2$ increases. This is because, the absolute gap is less than or equal to $FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2}) - FO_{P1}(\overline{\alpha}_2) = \sum_j h_j(r^*_j - (L_j \sum_i \mu_i Y^*_{ij} + z_{\overline{\alpha}_2}\sqrt{L_j \sum_i \sigma_i^2 Y^*_{ij}}))$. From table B.2 (appendix B.4) the maximum relative gap is 2.7%.

Regarding the number of DC installed we have the following comments derived from the computational experiments in table B.2:

- as expected, as the holding cost per unit and unit time increases, the number of DCs decreases;
- as expected, as the transport rate increases, the number of DCs increases;
- no effect of $\overline{\alpha}_1 - \overline{\alpha}_2$ on the number of DC is observed.

For all instances in table B.2, the total cost of the location-inventory model with differentiated service levels using critical level policy is less than the total cost induced by the global round-up policy. We measure the relative benefit induced by the critical level policy regarding global round-up policy as $\triangle cost(\%) = 100 \times (FO_{P1}(\overline{\alpha}_1) - FO^*_{P0})/FO_{P1}(\overline{\alpha}_1)$. Figure 5.2 show for each data set, how the relative benefit induced by the critical level policy change as the holding cost per unit and unit time (h_j) and the preset service level of the low priority class ($\overline{\alpha}_2$) change.

From figure 5.2, we can see that benefit induced by the critical level policy increase when the holding cost per unit and unit time and $\overline{\alpha}_1 - \overline{\alpha}_2$ increases. This is because, the absolute benefit is



(a) 49-node **(b)** 88-nodes

Figure 5.2: Relative benefit with $\overline{\alpha}_1 = 0.975$, $c_{ij} = 0.01$

greater than or equal to $FO_{P1}(\overline{\alpha}_1) - FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1}) = \sum_j h_j((L_j \sum_i \mu_i Y^*_{ij} + z_{\overline{\alpha}_1} \sqrt{L_j \sum_i \sigma_i^2 Y^*_{ij}}) - r^*_j)$. Therefore, we concluded that our location-inventory model with differentiable service levels using critical level policy is useful when the difference between the preset service levels for high and low priority class is high and / or the holding cost per unit and time unit is high. From table B.2 the maximum benefit induced by the critical level policy is 2.33%.

5.4 Conclusions

This chapter consider a location-inventory model for fast-moving items in which the distribution centers observe demand from two classes of customers, high and low priority. To provide differentiated service levels we assume, at each DC, a continuous review (Q, r, C) inventory policy. If a DC provides only one class of demand, the critical level policy becomes the traditional continuous review (Q, r) policy. In this chapter the service level is measured by service level type I.

We formulate the location-inventory model with differentiated service levels as an MINLP problem with chance constraints, corresponding to the service levels constraints, and nonlinear objective function. We propose optimally solve a relaxation of the location-inventory model with differentiated service levels which allows us to obtain good-quality bounds.

The computational results show that our proposed heuristic able to find good-quality solutions because for test set problems, the maximum optimality gap is 2.7%, a very good solution in itself, which provides us the configuration of the network, including location of the CD's, allocation of demands and the required stock everywhere.

The computational result also provides managerial insight: the benefit of using a critical level policy in the configuration of a distribution network is greater when the holding cost per unit and unit time is high, and / or when the difference between the preset service levels for high and low priority class is high.

There are a number of questions and issues left for future research. The first one, is to consider other policies to provide differentiated service levels in a distribution network, e.g., separate stock policy, single class allocation or round-up policy, so we can determine the best policy to provide differentiated service levels in a distribution network. The second one is related with the fact that our solution approach uses normally distributed demands. We believe that since the problem formulation is valid for any strictly increasing non-negative, similar solution approaches could be developed for other distributions in future research. Another possible extensions are: (i) consider other service levels measure, e.g., fill-rate; and (ii) use penalty cost as an alternative way of the service levels.

6 | On the Effect of Inventory Policies on Distribution Network Design with two Demand Classes

This chapter study several inventory policies in the design of a distribution network for fast-moving items able to provide differentiated service levels in terms of product availability for two demand classes (high and low priority). In particular, we consider the distribution network design problem when the *separate stock*, *single class allocation*, *global round-up*, *local round-up* and *critical level* inventory policies are used. We show how to formulate these problems as conic quadratic mixed-integer problems.

6.1 Model formulation

Consider the design of a distribution network consisting of an external supplier and a set of J, $j = 1, ..., |J|$, candidate sites for locating DCs which must supply a set I, $i = 1, ..., |I|$, of retailers. These retailers could be customers or markets, but for convenience we denote them as retailers in the rest of this chapter. We assume that the location of the external supplier, site candidates and retailers locations are known and that the supplier and DCs are uncapacitated. In this distribution network there are two categories of retailers or demand classes (high and low priority). The high priority retailers (class 1) require high service level and the low priority retailers (class 2) require lower service level. A retailer can be assigned to a single demand class, and we define $N_k = \{i \in I \mid i \text{ is class } k\}$, with $k = 1, 2$, as the set of retailers of class k. We also assume that the class of each retailer is known. The demands per unit time at each retailer are independent and normally distributed with mean $\mu_i > 0$ and variance $\sigma_i^2 > 0$. The problem is to determine the optimal number of DCs, their locations, the retailers assigned to each DC, how much inventory to keep at each of them and how to meet the preset service level for each demand class so as to minimize the total system cost.

In this chapter, each DC follow a continuous review (Q, r) policy with full-backorder and deterministic lead time that operate as follow. When the inventory position at DC j falls below a reorder level r_j, a replenishment order for Q_j units is placed and arrives $L_j > 0$ time units later. The service level level is measured by the service level type I. Let $\overline{\alpha}_k$ be the preset service level for class k, where $\overline{\alpha}_1 > \overline{\alpha}_2 > 0$. To provided differentiated service levels, we consider the *single class allocation*, *local separate stock* and *local round-up* policies.

As Daskin et al. (2002) and Shen et al. (2003) do, we assume the replenishment order Q_j is determined using an economic order quantity model (EOQ) and the steady-state backorders are negligible. The reorder point r_j depend the policy used to provide differentiated service levels. Hence, there are two types of decision variables in our model: the location variable, X_j; and the assignment variable, Y_{ij}. In particular, the last two variables are defined as:

$$X_j = \begin{cases} 1 & \text{if we locate a DC at candidate site } j \\ 0 & \sim \end{cases}$$

$$Y_{ij} = \begin{cases} 1 & \text{if demand at retailer } i \text{ is assigned to a DC at candidate site } j \\ 0 & \sim \end{cases}$$

Once we define the allocation variable Y_{ij}, we can characterize the demand at candidate DC j. Let $D_{kj}(\tau)$ be the total demand of class k during an interval of length τ at DC j and $D_j(\tau) = D_{1j}(\tau) + D_{2j}(\tau)$ be the total demand of both classes during an interval of length τ at DC j. As retailers, demand per unit time are independent and normally distributed, $D_{kj}(\tau)$ is also normally distributed with mean $\tau\mu_{kj}$ and variance $\tau\sigma^2_{kj}$, where $\mu_{kj} = \sum_{i\in N_k} \mu_i Y_{ij} \geq 0$ and $\sigma^2_{kj} = \sum_{i\in N_k} \sigma^2_i Y_{ij} \geq 0$. Furthermore, $D_j(\tau)$ is normally distributed with mean $\tau\mu_j$ and variance $\tau\sigma^2_j$, where $\mu_j = \mu_{1j} + \mu_{2j} = \sum_i \mu_i Y_{ij}$ and $\sigma^2_j = \sigma^2_{1j} + \sigma^2_{2j} = \sum_i \sigma^2_i Y_{ij}$. In appendix B.1, we present a glossary of terms.

6.1.1 Cost function

There is a fixed setup cost f_j of opening each distribution center. Each DC can serve more than one retailer, but each retailer should be only assigned to exactly one DC. The ordering cost from distribution center j is S_j. Linear transportation costs are incurred for shipment from the external supplier to distribution center j with unit cost a_j and from distribution center j to retailer i with unit cost d_{ij}. Using this notation, the average cost per unit time at DC j is:

$$AC_j(r_j, C_j, Y_{ij}) = S_j \frac{\sum_{i \in I} \mu_i Y_{ij}}{Q_j} + a_j \sum_{i \in I} \mu_i Y_{ij} + \sum_{i \in I} d_{ij} \mu_i Y_{ij} + h_j \left(\frac{Q_j}{2} + r_j - L_j \sum_{i \in I} \mu_i Y_{ij} \right). \quad (6.1)$$

The first term of equation (6.1) is the ordering cost per unit time. The second and third term are the supply and distribution costs per unit time respectively. As we assume negligible backorders, the fourth term is approximated the holding cost per unit time. Each distribution center determines the replenishment order Q_j using an EOQ model, i.e.,

$$Q_j = \sqrt{\frac{2S_j}{h_j} \sum_i \mu_i Y_{ij}} \, . \quad (5.7)$$

Then, replacing equation (5.7) into equation (6.1), the average cost per unit time at DC j is:

$$AC_j(r_j, C_j, Y_{ij}) = \sum_{i \in I} \hat{d}_{ij} Y_{ij} + k_j \sqrt{\sum_{i \in I} \mu_i Y_{ij}} + h_j \left(r_j - Lj \sum_{i \in I} \mu_i Y_{ij} \right) , \quad (6.2)$$

where $k_j = \sqrt{2h_j S_j}$ and $\hat{d}_{ij} = (a_j + d_{ij})\mu_i$.

In what follow, we formulate the inventory-location problems when *single class allocation*, *local separate stock* and *local round-up* policies are assumed to provide differentiated service levels. This three policies differ each of one on the reorder point formulation.

6.1.2 Inventory-Location Problem under local separate stock policy (SSP)

The local separate stock policy consider that each DC serves the demand assigned to it from a common stockpile and uses separate safety stocks for each class. The reorder point of the separate-stock policy is obtained from $r_j = \sum_i r_{kj}$, where r_{kj} is solution of $F_{D_{kj}(L_j)}(r_{kj}) = \alpha_k$. Under normally distributed demand:

$$\begin{aligned} r_j &= \mu_j L_j + \sum_{k\in K} z_{\alpha_k}\sigma_{kj}\sqrt{L_j} \\ &= L_j \sum_{i\in I}\mu_i Y_{ij} + \sum_{k\in K} z_{\alpha_k}\sqrt{L_j \sum_{i\in N_k}\sigma_i^2 Y_{ij}} \end{aligned} \tag{6.3}$$

where z_{α_k} is the inverse standard normal distribution for a preset service level α_k.

Replacing equation (6.3) in (6.2) and rearranging terms, we formulate an integrated location-inventory model under local separate stock policy as an INLP, denoted (**SSP**), as follows:

Problem (**SSP**):

$$\min_{X,Y} \quad \sum_{j\in J}\left\{ f_j X_j + \sum_{i\in I}\widehat{d}_{ij}Y_{ij} + \widehat{k}_j\sqrt{\sum_{i\in I}\mu_i Y_{ij}} + \widehat{\theta}_j \sum_{k\in K} z_{\alpha_k}\sqrt{\sum_{i\in N_k}\sigma_i^2 Y_{ij}} \right\} \tag{6.4}$$

$$\text{s.t:} \quad \sum_{j\in J} Y_{ij} = 1 \qquad \forall i\in I \tag{6.5}$$

$$Y_{ij} \le X_j \qquad \forall i\in I, j\in J \tag{6.6}$$

$$X_j,\ Y_{ij}\in\{0,1\} \qquad \forall i\in I, j\in J \tag{6.7}$$

where $\widehat{\theta}_j = h_j\sqrt{L_j}$. Constraints (6.5) establish that each customer is assigned exactly to one DC, (6.6) ensure that one customer can be assigned to location j only if a DC is installed there, and (6.7) are integrality constraints.

6.1.3 Inventory-location problem under single class allocation (SCA)

The *single class allocation* policy consider that each DC serves a single demand class. To formulate this policy, we define a new a new variable:

$$V_j^k = \begin{cases} 1 & \text{if a DC installed at } j \text{ serves class } k \\ 0 & \sim \end{cases} \tag{6.8}$$

Then, the integrated location-inventory model under single class allocation policy is formulated as an INLP, denoted (**SCA**), as follows:

Problem (**SCA**):

$$\min_{X,Y,V} \quad \sum_{j\in J}\left\{f_jX_j + \sum_{i\in I}\widehat{d}_{ij}Y_{ij} + \widehat{k}_j\sqrt{\sum_{i\in I}\mu_iY_{ij}} + \widehat{\theta}_j\sum_{k\in K}Z_{\alpha_k}\sqrt{\sum_{i\in N_k}\sigma_i^2Y_{ij}}\right\} \tag{6.4}$$

s.t: (6.5), (6.6), (6.7),

$$Y_{ij} \leq V_{kj} \qquad \forall i \in N_k, j \in J, k \in K \tag{6.9}$$

$$\sum_{k\in K}V_{kj} \leq X_j \qquad \forall\ j \in J \tag{6.10}$$

$$V_{kj} \in \{0,1\} \qquad \forall\ j \in J, k \in K. \tag{6.11}$$

Constraints (6.9) ensure the allocation of each customer to a DC that serves its class, (6.10) impose that each installed DC serves only one class, and (6.11) are the integrality constraints for the new variable V_{kj}.

It is easy to show that the problem (**SSP**) is a relaxation of (**SCA**). Therefore, the optimal solution of the problem (**SSP**) is a lower bound (LB) of problem (**SCA**).

6.1.4 Inventory-Location problem under local round-up policy (LRU)

The *local round-up policy* consider that each DC serves all demand assigned to it from a common stockpile and sets the safety stock as the maximum required between the sets of classes assigned to it. The reorder point of the local round-up policy at DC j is obtained from $F_{D_j(L_j)}(r_j) = \underset{k\in K}{max}\left\{\alpha_kV_j^k\right\}$ and under normally distributed demand:

$$\begin{aligned} r_j &= \mu_jL_j + \underset{k\in K}{max}\{z_{\alpha_k}V_{kj}\}\sigma_j\sqrt{L_j} \\ &= L_j\sum_{i\in I}\mu_iY_{ij} + \underset{k\in K}{max}\left\{z_{\alpha_k}V_{kj}\right\}\sqrt{L_j\sum_{i\in I}\sigma_i^2Y_{ij}} \end{aligned} \tag{6.12}$$

Replacing equation (6.12) in (6.2), define $Z_j = \underset{k\in K}{max}\left\{z_{\alpha_k}V_{kj}\right\}$ and rearranging terms, we formulate an location-inventory model under local round-up policy, denoted (**LRU**), as follows:

Problem (**LRU**):

$$\min_{X,Y,V,Z} \quad \sum_{j\in J}\left\{f_jX_j + \sum_{i\in I}\widehat{d}_{ij}Y_{ij} + \widehat{k}_j\sqrt{\sum_{i\in I}\mu_iY_{ij}} + \widehat{\theta}_jZ_j\sqrt{\sum_{i\in I}\sigma_i^2Y_{ij}}\right\} \tag{6.13}$$

s.t: (6.5), (6.6), (6.7), (6.9), (6.11)

$$Z_j \geq z_{\alpha_k}V_{kj} \qquad \forall j \in J, k \in K \tag{6.14}$$

$$Z_j \geq 0 \qquad \forall j \in J \tag{6.15}$$

Constraints (6.14) stipulate that Z_j is equal to the maximum of the inverse normal distribution for the service levels of the demand that the DC j serves and (6.15) are the non-negativity constraints for Z_j. This variable is greater or equal to zero, because we assume $\alpha_2 \geq 0.5$.

6.2 A Conic Quadratic MIP Formulation

In this section we show how to reformulate **SSP**, **SCA**, **LRU**, as a conic quadratic mixed-integer program (CQMIP). The advantage of the CQMIP formulation is that it can be solved directly using standard optimization software packages such as CPLEX or Mosek.

The square root term in the objective function of problems **SSP**, **SCA**, **LRU**, can give rise to difficulties in the optimization procedure. When the DC j is not selected, the square root terms would take a value of 0, which leads to unbounded gradients in the NLP optimization and hence numerical difficulties. Thus, we reformulate the problems **SSP**, **SCA**, **LRU** in order to eliminate the square root terms. We first introduce three sets of nonnegative continuous variables, $H1_j$, $H2_j$, $H3_{kj}$, to represent the square root terms in (6.4) and (6.13) :

$$H1_j^2 = \sum_{i\in I}\mu_iY_{ij}^2, \quad \forall j \in J \tag{6.16}$$

$$H2_j^2 = \sum_{i\in I}(\sigma_iY_{ij})^2, \quad \forall j \in J \tag{6.17}$$

$$H3_{kj}^2 = \sum_{i\in N_k}(\sigma_iY_{ij})^2, \quad \forall j \in J, k \in K \tag{6.18}$$

$$H1_j \geq 0, \quad \forall j \in J \tag{6.19}$$

$$H2_j \geq 0, \quad \forall j \in J \tag{6.20}$$

$$H3_{kj} \geq 0, \quad \forall j \in J, k \in K \tag{6.21}$$

Because the nonnegative variables $H1_j$, $H2_j$ and $H3_{kj}$ are introduced in the objective function of **SSP**, **SCA**, **LRU**, with positive coefficients, and this problems are minimization problem, the equations (6.16), (6.17) and (6.18) can be further relaxed as the following inequalities:

$$H1_j^2 \geq \sum_{i \in I} \mu_i Y_{ij}^2, \quad \forall j \in J \tag{6.22}$$

$$H2_j^2 \geq \sum_{i \in I} (\sigma_i Y_{ij})^2, \quad \forall j \in J \tag{6.23}$$

$$H3_{kj}^2 \geq \sum_{i \in N_k} (\sigma_i Y_{ij})^2, \quad \forall j \in J, k \in K \tag{6.24}$$

Note that constraints (6.22), (6.23), (6.24) with constraint (6.19), (6.20) and (6.21) define second-order cone constraints. Thus, the reformulated problems **SSP**, **SCA**, **LRU**, can be expressed as the following MINLP problem with second-order cone constraints denoted as ($\mathbf{CQ_{SSP}}$, $\mathbf{CQ_{SCA}}$, $\mathbf{CQ_{LRU}}$) respectively:

Problem ($\mathbf{CQ_{SSP}}$):

$$\min_{X,Y,H1,H3} \quad \sum_{j \in J} \left\{ f_j X_j + \sum_{i \in I} \widehat{d}_{ij} Y_{ij} + \widehat{k}_j H1_j + \widehat{\theta}_j \sum_{k \in K} z_{\alpha_k} H3_{kj} \right\} \tag{6.25}$$

s.t: (6.5), (6.6), (6.7), (6.22), (6.24), (6.19), (6.21)

Problem ($\mathbf{CQ_{SCA}}$):

$$\min_{X,Y,V,H1,H3} \quad \sum_{j \in J} \left\{ f_j X_j + \sum_{i \in I} \widehat{d}_{ij} Y_{ij} + \widehat{k}_j H1_j + \widehat{\theta}_j \sum_{k \in K} z_{\alpha_k} H3_{kj} \right\} \tag{6.25}$$

s.t: (6.5), (6.6), (6.7), (6.9), (6.10), (6.11), (6.22), (6.24), (6.19), (6.21)

Problem ($\mathbf{CQ_{SSP}}$) and ($\mathbf{CQ_{SCA}}$) can be trivially shown to be equivalent to problem (**SSP**) and (**SCA**) respectively but they have a linear objective function and second-order cone constraints (6.22) and (6.23) and (6.22) and (6.24) respectively. We can solve this problem using CPLEX 12.4., which handles second-order cone constraints in an efficient way.

Problem ($\mathbf{CQ_{LRU}}$):

$$\min_{X,Y,V,H1,H2} \quad \sum_{j\in J} \left\{ f_j X_j + \sum_{i\in I} \widehat{d}_{ij} Y_{ij} + \widehat{k}_j H1_j + \widehat{\theta}_j Z_j H2_j \right\} \tag{6.26}$$

s.t: (6.5), (6.6), (6.7), (6.9), (6.11), (6.14), (6.15)(6.22), (6.23), (6.19), (6.21)

Problem ($\mathbf{CQ_{LRU}}$) can be trivially shown to be equivalent to problem (**LRU**) but it has a non-linear objective function and second-order cone constraints (6.22) and (6.23). In what follow, we propose a solution approach for the ($\mathbf{CQ_{LRU}}$) problem that deal with the non-linear term on (6.26).

6.3 Solution Characterization for local round-up policy

Due to the non-linearity in the last term on (6.26), we need an special treatment to solve ($\mathbf{CQ_{LRU}}$). To achieve this, we relax (6.5) and add it to objective function with lagrangian multipliers λ_i , $\forall i \in I$. Then, the problem is separable into $|J|$ subproblems with the following structure:

Problem ($\mathbf{SP_{CQLRU}(j)}$):

$$\min_{X,Y,V,H1,H2} \quad f_j X_j + \sum_i (\widehat{d}_{ij} - \lambda_i) Y_{ij} + \widehat{k}_j H1_j + \widehat{\theta}_j Z_j H2_j \tag{6.27}$$

$$\text{s.t:} \quad Y_{ij} \le X_j \qquad \forall i \in I \tag{6.28}$$

$$Z_j \ge z_{\alpha_k} V_{kj} \qquad \forall k \in K \tag{6.29}$$

$$\sum_{i\in I} \sigma_i^2 (Y_{ij})^2 \le (H2_j)^2 \tag{6.30}$$

$$\sum_{i\in I} \mu_i (Y_{ij})^2 \le (H1_j)^2 \tag{6.31}$$

$$X_j, Y_{ij} \in \{0,1\} \qquad \forall i \in I \tag{6.32}$$

$$Z_j, H1_j, H2_j \ge 0 \tag{6.33}$$

Note that the objective function of ($\mathbf{SP_{CQLRU}(j)}$) remains non-linear, due to that, it can not be solved directly using standard optimization software packages such CPLEX or Mosek. To solve ($\mathbf{SP_{CQLRU}(j)}$), we fix V_{kj} for all possible combinations, i.e., a DC can serve the high priority class, the low priority class, or both classes at once. Fixing V_{kj} transforms Z_j in a parameter, eliminating non-linear term in (6.27). We solve the subproblems with fixed V_{kj}, choosing the one with minimum optimal value. This procedure gives a lower bound (LB) for ($\mathbf{CQ_{LRU}}$), which is computed using the following expression:

$$LB^{(k)} = \sum_{j \in J} SP^{(k)}(j) + \sum_{i \in I} \lambda_i^{(k)}, \tag{6.34}$$

where $SP^k(j)$ represents the optimal value for problem ($\mathbf{SP_{CQLRU}(j)}$) in iteration k.

Overall the Lagrangian Relaxation Method solves the lagrangian lowerbounds for each value of the multipliers, computes upper bounds, and updates the Lagrangian multipliers. Here we describe how we obtain an upper bound for ($\mathbf{CQ_{LRU}}$). The Lagrangian multipliers (or dual variables) are updated using the sub-gradient method. The details of this method are described in algorithm 4 in C.1. The upper bound (UB) of ($\mathbf{CQ_{LRU}}$) is obtained taking the values of X_j and V_{kj} from the relaxation ($\mathbf{SP_{CQLRU}(j)}$), then assigning each customer i to an opened DC j, i.e., with $X_j = 1$, serving its class, and having the smallest value of $\widehat{d_{ij}}$. Thus Y_{ij} is obtained, allowing the evaluation of $H1_j$ and $H2_j$, and finally the computation of the objective value of ($\mathbf{CQ_{LRU}}$), which is the upper bound (UB) at iteration k. The procedure to obtain the Upper Bound of ($\mathbf{CQ_{LRU}}$) is shown in algorithm 3 in C.1.

6.4 Computational Study

In order to illustrate the performance of each policy when varying the parameters of the distribution network we carried out computational experiments for instances with 49 and 88 nodes from Daskin (2011). We generated several test problems for each data set in which we compare the separate stock, single class allocation, global round-up, local round-up and critical level inventory policies. We denote these instances as *test sets*. In all cases, each retailer location is also a candidate DC location, i.e., there are as many candidate DC locations as retailer locations for each instance.

The test problems were generated around a base case with the following parameters: service level requirements $\alpha_1 = 0.975$ and $\alpha_2 = 0.75$; cost per unit to ship between retailer i and candidate DC site j, d_{ij} equal to the distance between retailer i and candidate DC j multiplied by a transport rate $c_{ij} = 0.01$, $\forall i, j$; and holding cost per unit and unit time at candidate DC site j, $h_j = 0.25$, $\forall j \in J$. Furthermore, all the test problems used the following common criterion and parameters: demand per unit time at each retailer is normally distributed with mean $\mu_i = U[10, 50]$ and coefficient of variation $CV_i = U[0.1, 0.4]$; class of the retail i, $n_i = \{1, 2\}$ with discrete uniform distribution; fixed (per unit time) cost of locating a DC at candidate site j, $f_j = U[200, 300]$; cost per unit to ship between external supplier and candidate DC site j, $a_j = 0.5$, $\forall j \in J$; ordering cost from candidate DC site j, $S_j = 1000, \forall j \in J$; and lead time, $L_j = \{2, 3, 4\}$ with discrete uniform distribution.

Problems $\mathbf{CQ_{SSP}}$ and $\mathbf{CQ_{SCA}}$ and the Lagrangian relaxation, to solve $\mathbf{CQ_{LRP}}$, were modeled with AMPL and solves with CPLEX 12.4. The location-inventory model using global round-up and critical level policies were also modeled with AMPL and solved with CPLEX 12.4, following the formulations presented in chapter 5. The time limit was set for 10800 seconds. All test were carried on a PC with Intel Core i7 2.3 GHz processor and 16 GB RAM.

We solved 60 problems, 30 for each data set. In each problem, we changed parameters relative to the base case. In particular the modified parameters were: the preset service level for low priority class (β_2), the holding cost per unit and unit time (h_j), and the transport rate (c_{ij}). Table C.1 (appendix C.2) shows the data-set used, the parameter modified from the base case; results of location-inventory model with differentiated service levels using critical level policy; results of the global round-up policy; and the relative difference between the total costs induced by critical level and global round-up policies.

For all instances in table C.1, the total cost of the location-inventory model with differentiated service levels using critical level policy is less than the total cost induced by the global round-up, local round-up, single class allocation and separate stock policies. Let Z_{LRU}, Z_{SCA}, Z_{SSP} be the objective function of problems ($\mathbf{CQ_{LRU}}$), ($\mathbf{CQ_{SCA}}$) and ($\mathbf{CQ_{LRU}}$) respectively; and Z_{GRU} and Z_{CLP} objective function of the location-inventory problem using global round-up and critical level policies respectively developed and solved in chapter 5. We measure the relative benefit induced by the critical level policy regarding the others policies as $benefit(\%) = 100 \times (Z_P - Z_{CLP})/Z_{CLP}$, where Z_P is the objective function of the others policies. Table 6.1 show for each data set, the minimum, maximum and average relative benefit induced by the critical level policy with respect the others policies.

	Data set					
	49-nodes			88-nodes		
Benefit (%)	min	average	max	min	average	max
CLP vs GRU	0.02	0.86	2.22	0.02	0.82	2.38
CLP vs LRU	0.02	0.82	2.24	0.13	0.75	2.38
CLP vs SCA	7.46	12.87	15.40	16.74	18.12	21.53
CLP vs SSP	0.01	0.62	1.64	0.01	0.69	1.80

Table 6.1: Benefit of the critical level policy vs other policies

Table 6.1 shows that the greatest benefit is obtained regarding single class allocation policy. For both data set, the maximum benefit has the same behavior: (i) critical policy induces maximum benefit level relative to global round-up and local round-up when the holding cost per unit and unit time is maximum ($h_j = 1.25$) and the preset service level of the low priority class is minimum ($\beta_2 = 0.55$); (ii) critical policy induces maximum benefit level relative to single class allocation when the holding cost per unit and unit time is maximum and transport rate is minimum, and (iii) critical policy induces maximum benefit level relative to separate stock when the holding cost per unit and unit time is maximum and the preset service level of the low priority class is maximum.

6.5 Conclusions

We study the effect of using global round-up, local round-up, single class allocation, separate stock and critical level policies in the problem of designing a distribution network able to provide differentiated service levels for two demand classes. We formulate a location-inventory using local round-up, single class allocation and separate stock policies as an INLP where the non-linearity is induced by the inventory cost. We show, in this chapter, how to formulate these problems as conic quadratic mixed-integer problems and in the special case of single class allocation we propose a Lagrangian relaxation over its conic quadratic mixed-integer to solve it.

For the instances solved in the computational experiments, the best policy was never the one that assigns to each distribution center a unique demand class, i.e., the single class allocation is the worst policy performance in terms of total cost.

There are a number of questions and issues left for future research. The first one is to determine analytically the conditions under which a policy is better than the others, in particular, to generalize what we observed in our computational experiments in the sense that assigning a unique class to each DC is not the best policy. The second one is to use higher dimension network, e.g., for instances of 49 and 88 nodes by Daskin (2011). Another stream of future work is to broaden the measure of service level, e.g., fill rate, or to use a critical level policy to provide differentiated service levels.

7 | Illustrative Example for an Industrial Application

To illustrate the industrial applicability of the location-inventory model with differentiated service levels developed in chapters 5 and 6, we consider a illustrative example of a company that manufactures products derived from fruits and vegetables. The supply chain consisting of one plant, 38 potencial DCs, and 38 customers.

We will analyze five inventory policies to provided differentiated service levels for two demand classes. In particular, we consider the distribution network design problem when the *critical level*, *separate stock*, *single class allocation*, *global round-up* and *local round-up* policies are used. The location-inventory model that use a critical level inventory policy to provide differentiated service levels was developed in chapter 5, while the location-inventory models that use separate stock, single class allocation, global round-up and round-up policies was developed in chapter 6. Therefore, for the Illustrative example, we also determine the best inventory policy implemented in the distribution network to provide differentiated service levels.

7.1 Parameters: product, customers, potencial DCs and costs

Consider the case of a company that manufactures products derived from fruits and vegetables which requires determining the number of distribution centers (DC) to locate in Santiago (Chile), where to locate, what kinds of customers should be assigned to each DC, how much inventory to keep each of them, and how to meet the required service level of their customer. The supply chain consisting of one plant, 38 potencial DCs, and 38 customers. The production plant is located 200 *km* south of Santiago (Chile). The company segments its customers by volume of annual demand. Thus, customers who demand more than the average annual demand are classified as high priority and its preset service level is $\beta_1 = 0.98$. Customers who require less than the average annual demand are classified as low priority and its preset service level is $\beta_2 = 0.70$.

The products manufactured by the company are derivative of fruits and vegetables, with holding cost per unit and unit time at candidate DC site j, $h_j = 0.005(US\$/Kg-day)$, $\forall j \in J$. The location, class, demand, and coefficient of variance for each customers is show in table D.1. From table D.1, note that 24% of customers are class 1 and they demand 72% of daily demand.

Ordering cost from candidate DC site j, $S_j = 250(US\$/order), \forall j \in J$; cost per unit to ship between plant and candidate DC site j, $a_j = 0.0069(US\$/kg)$, $\forall j \in J$; and lead time, $L_j = 4(day)$, $\forall j \in J$. The location and fixed cost to the 40 potential DCs is show in table D.1. For class 1 customers the company uses medium goods vehicles and for class 2 customers the company uses light goods vehicle. Each vehicle uses a driver and an assistant. The unit cost of transport $(US\$/Kg)$ between candidate DC j and retail i, d_{ij}, is calculated as the fixed cost of loading and unloading (including labor and depreciation), plus variable cost that depends on the distance between candidate DC j and retail i (including labor, fuel and depreciation). Then, unit cost of transport are $d_{ij} = 0.0025 + 0.00012\ \theta_{ij}$ and $d_{ij} = 0.0021 + 0.00015\ \theta_{ij}$ for class 1 and 2 respectively, where θ_{ij} is the Euclidean distance between candidate DC j and retail i.

7.2 Network configuration using several inventory control policies

We analyze the location-inventory problem with differentiated service levels using five inventory control policies. Table 7.1 shows the inventory policy, number of opened DC, the objective function and the cost components (FC: fixed cost; OC: ordering cost; SC: supply cost; CD: distribution cost; HC: holding cost).

	Result			Cost Component (%)				
Policy	♯ DC	FO	Time (s)	FC	OC	SC	DC	HC
Critical Level	1	807.2	0.140	0.2	0.2	0.2	0.1	0.3
Single Class Allocation	2	1122.1	13.361	0.3	0.2	0.1	0.1	0.2
Separate stock policy	1	829.5	0.149	0.2	0.1	0.2	0.1	0.3
Global round-up	1	826.7	0.149	0.2	0.1	0.2	0.1	0.3
Local Round - up	1	826.7	41.297	0.2	0.1	0.2	0.1	0.3

Table 7.1: Illustrative example: results.

Table 7.1 indicates that the lower cost configuration is achieved with the critical level policy with a saving of 39%, 2.8%, 2.4% and 2.4% per day with respect to single class allocation, separate stock, global round-up and local round-up respectively. The resulting network configuration is the same for critical level, separate stock, global round-up and local round-up policies. Figure 7.1 show the network configuration.

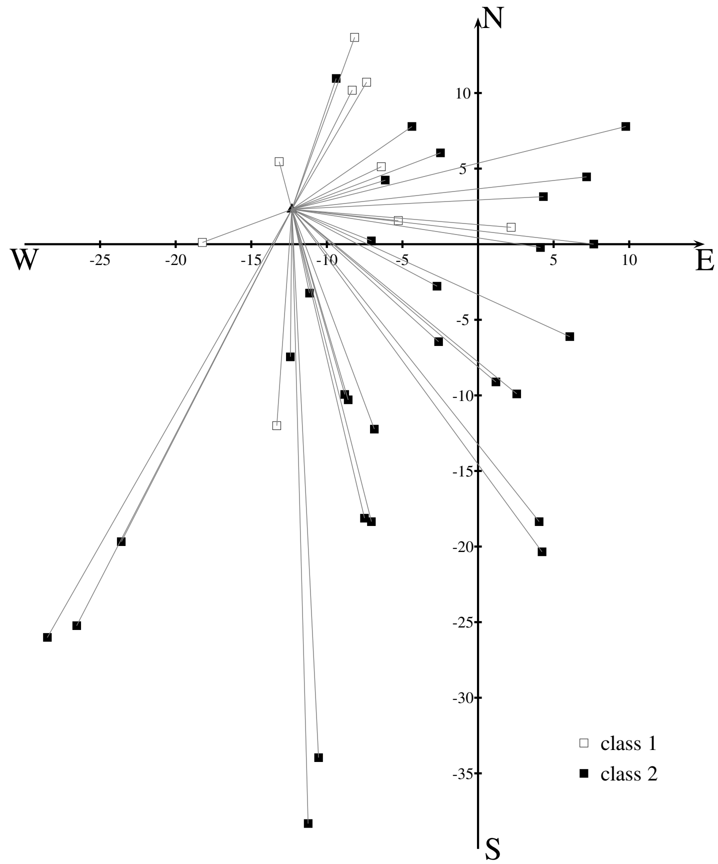


Figure 7.1: Network configuration using Critical Level policy : $FO^* = 807.2$

Single class allocation policy induces the highest cost because two DCs are installed on the network. Figure 7.2 show the network configuration for the single class allocation policy.

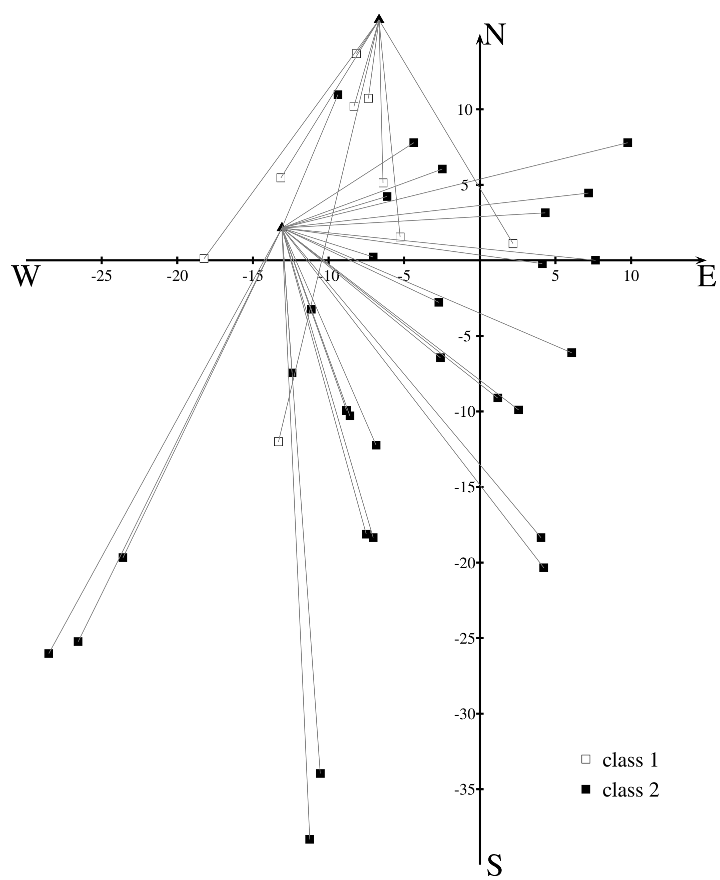


Figure 7.2: Network configuration using single class allocation policy : $FO^{*} = 1122.1$

8 | Conclusions

Location-inventory mathematical models have been known since 2000. Earlier models were limited by the assumption that each customer has identical variance-to-mean ratio. Furthermore, models that integrate inventory and location decisions, consider that the distribution network is dominated by continuous review (Q, r) policy, full backorder, deterministic lead time, normally distributed demand and service level type I, in which the same preset service level is provided to the whole network. Therefore, previous location-inventory models do not integrate the requirement of different demand classes in the optimal configuration of the distribution network.

The design of distribution networks able to provide differentiated service levels is particularly relevant for fast-moving items because in the last decades the distribution channels of fast-moving consumer goods (FMCG) have been concentrated on large retails chains, which demand a large quantity of items and are therefore in a position to request high service level in terms of product availability at the supplier's expense. In this thesis, we consider a supplier of fast moving items that serves several customers, including large retails chains, and is likely to face a stock-out. Given this situation, the supplier would likely prefer to meet the higher service level requested by the large retail chain to ensure a good relationship with the businesses that most impact the bottom line. This makes a natural situation where the supplier decides to meet demand with differentiated service levels and segment operationally their customers based on service levels. The simplest segmentation is to classify customers into two demand classes: (i) High-priority class that would correspond to large retail chains that require high levels of service and, (ii) Low-priority class that would represent small retailers that have to settle for a lower level of service.

This thesis consists of several small steps towards constructing an efficient supply chain management system for fast moving-items, which is also able to provide differentiated service levels to two demand classes (high and low priority). We propose five location-inventory models with differentiated service levels using (i) *single class allocation*, in which each DC serves a single demand class; (ii) *global round-up policy*, which set the service level of the entire distribution network based on a preset level corresponding to the highest priority class; (iii) *separate stock policy*, which impose that each DC serves the demand assigned to it from a common stockpile and uses separate safety stocks for each class; (iv) *local round-up policy*, which serves all demand assigned to it from a common stockpile and sets the safety stock as the maximum required between the sets of classes assigned to it; and (v) *critical level policy*, which rationing the inventory between different classes.

Except for the critical level policy, all the mechanisms for providing differentiated service levels operating under continuous review policy and strictly increasing non-negative demand may extend to distribution networks. In general, the complexity is not in its formulation, but in the solution method due to the non-linearity that inventories introduce in the the network configuration. For critical level policy it is first necessary to develop a theoretical formulation with continuous review and strictly increasing non-negative demand, and then integrate it into a location-inventory model, because previous work on critical level policy has only considered the case of slow-moving items.

The contributions in this thesis are:

- In chapter 3 (i) we present a new critical level inventory model and solution technique when demand is modeled through continuous distributions, (ii) we develop expressions for the service level type 1 under rationing and (iii) we provide exact expressions for the steady state backorders under rationing.

- In chapter 4 we model and solve a critical level model when demand is modeled through continuous distributions and backorders are penalized with differentiated costs. Furthermore, we develop approximate expressions for backorders of each class and we demonstrate its convexity.

- In chapter 5 (i) we address for the first time the modelling and solution of a supply chain design problem of fast moving items that considers the ability of the distribution network to provide and fulfill different service levels in term of product availability, (ii) we demonstrate that under no demand for one of the classes, the (Q, r, C) policy is equivalent to the traditional (Q, r) policy and (iii) the service level constraints, under rationing, remain valid under no demand for one of the classes.

- In chapter 6 we model and solve three alternatives ways to provide differentiated service levels in a distribution network that can extend the network design alternatives able to provide differentiated service levels.

Regarding the use of critical level policy in a single-echelon inventory system for fast-moving items we conclude that:

- the average savings induced by the critical level policy are greater with respect to separate stock and round up policy,
- critical level policy leads to significant savings with respect to round-up when class 2 dominates on mean and variance, the ordering cost is minimal, holding cost per unit and unit time is maximum and difference between preset service levels (or shortage costs) is maximum,
- critical level policy leads to significant savings with respect to separate stock when class 2 dominates on mean, class 1 dominates on variance, the ordering cost is minimal, holding cost per unit and unit time is maximum and difference between preset service levels is maximum (or difference between the shortage costs is minimal),
- the cost of increasing the service level of the high priority class is significantly greater than the cost of increasing the service level of the low priority class.

Regarding the distribution network that include two class demand and fast-moving items, we conclude that:

- the benefit of using a critical level policy in the configuration of a distribution network with respect to global round-up and local round-up when tthe holding cost per unit and unit time is high, and / or when the difference between the preset service levels for high and low priority class is high;
- the benefit of using a critical level policy in the configuration of a distribution network with respect to single class allocation when the holding cost per unit and unit time is maximum and transport rate is minimum
- the benefit of using a critical level policy in the configuration of a distribution network with respect to separate stock when the holding cost per unit and unit time is maximum and the preset service level of the low priority class is maximum.
- the best policy was never the one that assigns to each distribution center a unique demand class, i.e., the single class allocation is the worst policy performance in terms of total cost.

Bibliography

Arslan, H., Graves, S. C., and Roemer, T. (2007). A single-product inventory model for multiple demand classes. *Management Science*, 53:1486–1500.

Atamtürk, A., Berenguer, G., and Shen, Z.-J. (2012). A conic integer programming approach to stochastic joint location-inventory problems. *Operations Research*, 60(2):366–381.

Axsäter, S. (2007). *Inventory control*, volume 90. Springer.

Axsäter, S. (1996). Using the deterministic eoq formula in stochastic inventory control. *Management Science*, 42:830–834.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

Browne, S. and Zipkin, P. (1991). Inventory models with continuous, stochastic demands. *The Annals of Applied Probability*, pages 419–435.

Bulut, Ö. and Fadiloğlu, M. M. (2011). Production control and stock rationing for a make-to-stock system with parallel production channels. *IIE Transactions*, 43(6):432–450.

Daskin, M. S. (2011). *Network and discrete location: models, algorithms, and applications*. John Wiley & Sons.

Daskin, M. S., Coullard, C. R., and Shen, Z.-J. M. (2002). An inventory-location model: Formulation, solution algorithm and computational results. *Annals of Operations Research*, 110(1-4):83–106.

De Vericourt, F., Karaesmen, F., and Dallery, Y. (2000). Dynamic scheduling in a make-to-stock system: A partial characterization of optimal policies. *Operations Research*, 48(5):811–819.

De Vericourt, F., Karaesmen, F., and Dallery, Y. (2002). Optimal stock allocation for a capacitated supply system. *Management Science*, 48(11):1486–1501.

Dekker, R., Hill, R. M., Kleijn, M. J., and Teunter, R. H. (2002). On the (s-1,s) lost sales inventory model with priority demand classes. *Naval Research Logistics*, 49:593–610.

Dekker, R., Kleijn, M. J., and De Rooij, P. (1998). A spare parts stocking policy based on equipment criticality. *International Journal of Production Economics*, 56:69–77.

Deshpande, V., Cohen, M. A., and Donohue, K. (2003). A threshold inventory rationing policy for service-differentiated demand classes. *Management Science*, 49:683–703.

Eppen, G. (1979). Effect of centralization on expected cost in a multi-location newsboy problem. *Magement Science*, 25:498–501.

Erlenkotter, D. (1978). A dual-based procedure for uncapacitated facility location. *Oper Res*, 26:992–1009.

Fadiloglu, M. and Bulut, O. (2010). A dynamic rationing for continuous-review inventory systems. *European Journal of Operational Research*, 202:675–685.

Fadıloğlu, M. M. and Bulut, Ö. (2010). An embedded markov chain approach to stock rationing. *Operations Research Letters*, 38(6):510–515.

Ha, A. Y. (1997a). Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, 43(8):1093–1103.

Ha, A. Y. (1997b). Stock-rationing policy for a make-to-stock production system with two priority classes and backordering. *Naval Research Logistics (NRL)*, 44(5):457–472.

Ha, A. Y. (2000). Stock rationing in an m/ek/1 make-to-stock queue. *Management Science*, 46(1):77–87.

Isotupa, K. S. (2006). An (s,q) markovian inventory system with lost sales and two demand classes. *Mathematical and Computer Modelling*, 43:687–694.

Kleijn, M. and Dekker, R. (1999). An overview of inventory systems with several demand classes. In Speranza, M. and Stähly, P., editors, *New Trends in Distribution Logistics*, volume 480 of *Lecture Notes in Economics and Mathematical Systems*, pages 253–265. Springer Berlin Heidelberg.

Meerschaert, M. M. and Scheffler, H.-P. (2008). Triangular array limits for continuous time random walks. *Stochastic Processes and their Applications*, 118(9):1606 – 1633.

Melchiors, P., Dekker, R., and Kleijn, M. (2000). Inventory rationing in an (s,q) inventory model with lost sales and two demand classes. *Journal of Operational Research Society*, 51:111–122.

Miranda, P. A. and Garrido, R. A. (2004). Incorporating inventory control decisions into a strategic distribution network design model with stochastic demand. *Transportation Research Part E: Logistics and Transportation Review*, 40(3):183–207.

Möllering, K. and Thonemann, U. (2010). An optimal constant level rationing policy under service level constraints. *OR Spectrum*, 32(2):319–341.

Nahmias, S. and Demmy, W. S. (1981). Operating characteristics of an inventory system with rationing. *Management Science*, 27:1236–1245.

Ozsen, L., Coullard, C. R., and Daskin, M. S. (2008). Capacitated warehouse location model with risk pooling. *Naval Research Logistics (NRL)*, 55(4):295–312.

Park, C. and Padgett, W. J. (2005). Accelerated degradation models for failure based on geometric brownian motion and gamma processes. *Lifetime Data Analysis*, 11:511–527.

Peterson, R. and Silver, E. A. (1979). *Decision systems for inventory management and production planning*. Wiley New York.

Piplani, R. and Liu, S. (2014). Comparison of capacity rationing policies for a make-to-order production system with two customer classes. *International Journal of Industrial and Systems Engineering*, 16(2):223–237.

Ramaekers, K. and Janssens, G. K. (2008). On the choice of a demand distribution for inventory management models. *European Journal of Industrial Engineering*, 2(4):479–491.

Sadjadi, S. J., Makui, A., Dehghani, E., and Pourmohammad, M. (2015). Applying queuing approach for a stochastic location-inventory problem with two different mean inventory considerations. *Applied Mathematical Modelling*.

Serfozo, R. and Stidham, S. (1978). Semi-stationary clearing processes. *Stochastic Processes and their Applications*, 6(2):165–178.

Shahabi, M., Unnikrishnan, A., Jafari-Shirazi, E., and Boyles, S. D. (2014). A three level location-inventory problem with correlated demand. *Transportation Research Part B: Methodological*, 69:1–18.

Shen, Z.-J. M. (2005). A multi-commodity supply chain design problem. *IIE Transactions*, 37(8):753–762.

Shen, Z.-J. M., Coullard, C., and Daskin, M. S. (2003). A joint location-inventory model. *Transportation Science*, 37(1):40–55.

Shen, Z.-J. M. and Daskin, M. S. (2005). Trade-offs between customer service and cost in integrated supply chain design. *Manufacturing & service operations management*, 7(3):188–207.

Shu, J., Teo, C.-P., and Shen, Z.-J. M. (2005). Stochastic transportation-inventory network design problem. *Operations Research*, 53(1):48–60.

Snyder, L. V., Daskin, M. S., and Teo, C.-P. (2007). The stochastic location model with risk pooling. *European Journal of Operational Research*, 179(3):1221–1238.

Teunter, R. H. and Haneveld, W. K. K. (2008). Dynamic inventory rationing strategies for inventory systems with two demand classes, poisson demand and backordering. *European Journal of Operational Research*, 190(1):156 – 178.

Wang, D., Tang, O., and Huo, J. (2013a). A heuristic for rationing inventory in two demand classes with backlog costs and a service constraint. *Computers & Operations Research*, 40(12):2826–2835.

Wang, Y., Zhang, S., and Sun, L. (2013b). Anticipated rationing policy for two demand classes under service level constraints. *Computers and Industrial Enginnering*, 23:1200–1212.

You, F. and Grossmann, I. E. (2008). Mixed-integer nonlinear programming models and algorithms for large-scale supply chain design with stochastic inventory management. *Industrial & Engineering Chemistry Research*, 47(20):7802–7817.

Zheng, Y.-S. (1992). On properties of stochastic inventory systems. *Management science*, 38(1):87–103.

Zipkin, P. (1986). Inventory service-level measures: convexity and approximation. *Management Science*, 32(8):975–981.

A | Appendix Chapter 3

A.1 Proof of proposition 2

Lemma A.1.1. *Let X, Y be two univariate continuous random variables, where Y has positive support. Then, for any C we have*

$$\mathbb{P}(X + Y > C) \geq \mathbb{P}(X > C) \, .$$

Proof. We note that the set of realizations $\{\omega \mid X(\omega) > C\} \subset \{\omega \mid X(\omega) + Y(\omega) > C\}$, which gives the inequality $\mathbb{P}(X + Y > C) \geq \mathbb{P}(X > C)$. □ □

Given the lemma A.1.1, the demonstration of proposition 2 is:

Proof. Let $\tau_{R,D}^{r-C} = \min\{\tau_{H,D}^{r-C}, L\}$ be the *time to rationing*, which corresponds to the amount of time that elapses from the moment an order is placed until the critical level C is reached if this event occurs during the lead time. If the hitting time $\tau_{H,D}^{r-C}$ does not occur during lead time then the time to rationing is defined as $\tau_{R,D}^{r-C} = L$. In this case, rationing coincides with the reception of the replenishment batch, and therefore, to be precise, rationing is not produced.

Given a $k > 0$ we have that, for every demand realization ω, the hitting time satisfies $\tau_{H,D}^{r-C}(\omega) < \tau_{H,D}^{r+k-C}(\omega)$. This is because exactly k additional units of demand are necessary to reach C, and the demand is a strictly increasing non-negative demand. This implies that for any $k > 0$ we have that

$$\begin{aligned} \tau_{R,D}^{r-C}(\omega) < \tau_{R,D}^{r+k-C}(\omega) \qquad & \forall \omega \text{ s.t. } \tau_{H,D}^{r-C}(\omega) < L, \\ L = \tau_{R,D}^{r-C}(\omega) = \tau_{R,D}^{r+k-C}(\omega) \qquad & \forall \omega \text{ s.t. } \tau_{H,D}^{r-C}(\omega) \geq L \, . \end{aligned}$$

From these relations we have that $\tau_{R,D}^{r+k-C} - \tau_{R,D}^{r-C} \geq 0$ with probability 1, which combined with the assumptions on the demand gives us $D_1(L - \tau_{R,D}^{r-C}) = D_1(L - \tau_{R,D}^{r+k-C}) + D_1(\tau_{R,D}^{r+k-C} - \tau_{R,D}^{r-C})$, where the last term is a positive support random variable when $\tau_{H,D}^{r-C} < L$. We therefore have

$$\begin{aligned}
\alpha_1(r, C) &= \mathbb{P}(D_1(L - \tau_{R,D}^{r-C}) \leq C) \\
&= \mathbb{P}(D_1(L - \tau_{R,D}^{r-C}) \leq C \mid \tau_{H,D}^{r-C} \geq L)\mathbb{P}(\tau_{H,D}^{r-C} \geq L) + \mathbb{P}(D_1(L - \tau_{R,D}^{r-C}) \leq C \mid \tau_{H,D}^{r-C} < L)\mathbb{P}(\tau_{H,D}^{r-C} < L) \\
&= \mathbb{P}(D_1(L - \tau_{R,D}^{r+k-C}) \leq C \mid \tau_{H,D}^{r-C} \geq L)\mathbb{P}(\tau_{H,D}^{r-C} \geq L) \\
&\quad + \mathbb{P}(D_1(L - \tau_{R,D}^{r+k-C}) + D_1(\tau_{R,D}^{r+k-C} - \tau_{R,D}^{r-C}) \leq C \mid \tau_{H,D}^{r-C} < L)\mathbb{P}(\tau_{H,D}^{r-C} < L) \\
&\leq \mathbb{P}(D_1(L - \tau_{R,D}^{r+k-C}) \leq C \mid \tau_{H,D}^{r-C} \geq L)\mathbb{P}(\tau_{H,D}^{r-C} \geq L) + \mathbb{P}(D_1(L - \tau_{R,D}^{r+k-C}) \leq C \mid \tau_{H,D}^{r-C} < L)\mathbb{P}(\tau_{H,D}^{r-C} < L) \\
&= \mathbb{P}(D_1(L - \tau_{R,D}^{r+k-C}) \leq C) = \alpha_1(r + k, C) \,.
\end{aligned}$$

Here the inequality uses lemma A.1.1 with $X = D_1(L - \tau_{R,D}^{r+k-C})$ and $Y = D_1(\tau_{R,D}^{r+k-C} - \tau_{R,D}^{r-C})$.

We repeat the argument to show the tendency of $\alpha_1(r, C)$ with respect to C. Given any $k > 0$ we have that $\tau_{H,D}^{r-(C+k)}(\omega) < \tau_{H,D}^{r-C}(\omega)$ for any demand realization ω. Similarly, for any $k > 0$, we now have

$$\begin{aligned}
\tau_{R,D}^{r-(C+k)}(\omega) < \tau_{R,D}^{r-C}(\omega) \qquad &\forall \omega \text{ s.t. } \tau_{H,D}^{r-(C+k)}(\omega) < L \\
L = \tau_{R,D}^{r-(C+k)}(\omega) = \tau_{R,D}^{r-C}(\omega) \qquad &\forall \omega \text{ s.t. } \tau_{H,D}^{r-(C+k)}(\omega) \geq L \,.
\end{aligned}$$

The demand can be now separated $D_1(L - \tau_{R,D}^{r-(C+k)}) = D_1(L - \tau_{R,D}^{r-C}) + D_1(\tau_{R,D}^{r-C} - \tau_{R,D}^{r-(C+k)})$, where for every demand realization ω this last term satisfies $D_1(\tau_{R,D}^{r-C} - \tau_{R,D}^{r-(C+k)})(\omega) \leq k$. This because $\tau_{R,D}^{r-C}(\omega) - \tau_{R,D}^{r-(C+k)}(\omega) \leq \tau_{H,D}^{r-C}(\omega) - \tau_{H,D}^{r-(C+k)}(\omega)$ and $D_1(\tau_{H,D}^{r-C} - \tau_{H,D}^{r-(C+k)}) \leq D(\tau_{H,D}^{r-C} - \tau_{H,D}^{r-(C+k)}) = k$ by definition of hitting time. This gives

$$\begin{aligned}
\alpha_1(r, C) &= \mathbb{P}(D_1(L - \tau_{R,D}^{r-C}) \leq C) = \mathbb{P}(D_1(L - \tau_{R,D}^{r-C}) \leq C \mid \tau_{H,D}^{r-(C+k)} \geq L)\mathbb{P}(\tau_{H,D}^{r-(C+k)} \geq L) \\
&\quad + \mathbb{P}(D_1(L - \tau_{R,D}^{r-C}) \leq C \mid \tau_{H,D}^{r-(C+k)} < L)\mathbb{P}(\tau_{H,D}^{r-(C+k)} < L) \\
&= \mathbb{P}(D_1(L - \tau_{R,D}^{r-(C+k)}) \leq C + k \mid \tau_{H,D}^{r-(C+k)} \geq L)\mathbb{P}(\tau_{H,D}^{r-(C+k)} \geq L) \\
&\quad + \mathbb{P}(D_1(L - \tau_{R,D}^{r-(C+k)}) \leq C + D_1(\tau_{R,D}^{r-C} - \tau_{R,D}^{r-(C+k)}) \mid \tau_{H,D}^{r-(C+k)} < L)\mathbb{P}(\tau_{H,D}^{r-(C+k)} < L) \\
&\leq \mathbb{P}(D_1(L - \tau_{R,D}^{r-(C+k)}) \leq C + k \mid \tau_{H,D}^{r-(C+k)} \geq L)\mathbb{P}(\tau_{H,D}^{r-(C+k)} \geq L) \\
&\quad + \mathbb{P}(D_1(L - \tau_{R,D}^{r-(C+k)}) \leq C + k \mid \tau_{H,D}^{r-(C+k)} < L)\mathbb{P}(\tau_{H,D}^{r-(C+k)} < L) \\
&= \mathbb{P}(D_1(L - \tau_{R,D}^{r-(C+k)}) \leq C) = \alpha_1(r, C + k) \,.
\end{aligned}$$

Here we add a k in the first term of the second equality because $D_1(L - \tau_{R,D}^{r-(C+k)}) = D_1(0)$ when $\tau_{H,D}^{r-(C+k)} \geq L$, so that first probability equals 1. The inequality comes from the fact that $\mathbb{P}(D_1(\tau_{R,D}^{r-C} - \tau_{R,D}^{r-(C+k)}) \leq k) = 1$. □

A.2 Partial derivative of $\alpha_1(r, C)$ with respect to C

Here we give the expression of $\frac{\partial \alpha_1(r,C)}{\partial C}$ in the case when the demands for both classes are normally distributed and the density function of the hitting time $\tau_{H,D}^{r-C}$ is given by equation (3.19). We denote by $\bar{\varphi}_{\mu,\sigma^2}(x)$ the density function of a normal random variable with mean μ and variance σ^2. The partial derivative then can be expressed as:

$$\frac{\partial \alpha_1(r,C)}{\partial C} = \int_0^L \left(\frac{r - C + \mu\tau}{2\tau} - \frac{C + \mu_1(L-\tau)}{2(L-\tau)} \right) \bar{\varphi}_{\mu_1(L-\tau),\sigma_1^2(L-\tau)}(C) \bar{\varphi}_{\mu\tau,\sigma^2\tau}(r - C) d\tau \ . \tag{A.1}$$

B | Appendix Chapter 5

B.1 Glossary of terms

Sets	Definition
I	Set of retailers indexed by i
J	Set of candidate DC sites indexed by j
K	Set of class demand indexed by k, with $k = 1, 2$
N_k	Set of retailers that belong to the class k, with $k = 1, 2$
Parameters	
μ_i	mean demand per unit time at retailer i
σ_i	Standard deviation of demand per unit time at retailer i
$\overline{\alpha}_k$	preset service level for class k, with $\overline{\alpha}_1 > \overline{\alpha}_2$
f_j	Fixed (per unit time) cost of locating a DC at candidate site j
d_{ij}	Cost per unit to ship between retailer i and candidate DC site j
c_{ij}	transport rate between retailer i and candidate DC j
a_j	Cost per unit to ship between external supplier and candidate DC site j
S_j	Ordering cost from candidate DC site j
h_j	Holding cost per unit time at candidate DC site j
L_j	Constant replenishment lead time at candidate DC site j
n_i	Class of retail i
Variables	
X_j	1 if we locate a DC in candidate site j, and 0 otherwise
Y_{ij}	1 if retailer i is served by the DC at candidate site j, and 0 otherwise
r_j	reorder point at candidate DC site j
C_j	critical level at candidate DC site j
Variable functions	
$\mu_j = \sum_i \mu_i Y_{ij} \geq 0$	mean demand per unit time at candidate DC j
$\sigma_j = \sqrt{\sum_i \sigma_i^2 Y_{ij}} \geq 0$	standard deviation of demand per unit time at candidate DC j
$\mu_{kj} = \sum_{i \in N_k} \mu_i Y_{ij} \geq 0$	mean demand per unit time of class k at candidate DC j
$\sigma_{kj} = \sqrt{\sum_{i \in N_k} \sigma_i^2 Y_{ij}} \geq 0$	standard deviation of demand per unit time of class k at candidate DC j
$Q_j = \sqrt{\frac{2S_j}{h_j} \sum_i \mu_i Y_{ij}}$	replenishment order at candidate DC j

Table B.1: Glossary of terms

B.2 Proof of proposition 8

To prove that equations (5.1) and (5.2) are general expressions for the service level type I under rationing we must show that these equations remain valid in the absence of one of the two class demand. In order to simplify the notation, we suppress the subscript j.

Proof. The conditions to fully meet the demand of class 1 in a replenishment cycle when the DC only provides class 1 is that demand of class 1 during the lead time is less or equal to the reorder point r, i.e., $D_1(L) \leq r$. Therefore, the service level provided to the high priority class when the DC only provides service to class 1 is:

$$\alpha_1(r, C) = \mathbb{P}(D_1(L) \leq r). \tag{B.1}$$

Equation (B.1) is equal to equation (5.1) when $D_2(\tau) = 0$ for any $\tau > 0$, because

$$\begin{aligned}
P(D_1(L) \leq r) =& \mathbb{P}(D_1(L) \leq r \mid D_1(L) \leq r - C))\mathbb{P}(D_1(L) \leq r - C) + \mathbb{P}(D_1(L) \leq r \mid D_1(L) > r - C))\mathbb{P}(D_1(L) > r - C) \\
=& \mathbb{P}(D_1(L) \leq r - C) + \mathbb{P}(D_1(L) \leq r \mid D_1(L) > r - C))\mathbb{P}(D_1(L) > r - C) \\
=& \mathbb{P}(D_1(L) \leq r - C) + \mathbb{P}(D_1(L) \leq r - C + C \mid \tau_{H,D_1}^{r-C} < L)\mathbb{P}(\tau_{H,D_1}^{r-C} < L) \\
=& \mathbb{P}(D_1(L) \leq r - C) + \mathbb{P}(D_1(L) \leq D_1(\tau_{H,D_1}^{r-C}) + C \mid \tau_{H,D_1}^{r-C} < L)\mathbb{P}(\tau_{H,D_1}^{r-C} < L) \\
=& \mathbb{P}(D_1(L) \leq r - C) + \mathbb{P}(D_1(L - \tau_{H,D_1}^{r-C}) \leq C \mid \tau_{H,D_1}^{r-C} < L)\mathbb{P}(\tau_{H,D_1}^{r-C} < L) \\
=& \mathbb{P}(D_1(L) \leq r - C) + \mathbb{P}(D_1(L - \tau_{H,D_1}^{r-C}) \leq C \ \cap \ \tau_{H,D_1}^{r-C} < L).
\end{aligned}$$

Furthermore, in subsection 5.2.2 we show that in the absence of demand for class 1, $\alpha_1(r^*, C^*) = 1 > \overline{\alpha}_1$, where (r^*, C^*) are the optimal reorder point and critical level respectively. Therefore, equation (5.1) is a general expression for the service level provided to the high priority class.

The conditions to fully meet the demand of class 2 in a replenishment cycle, when the DC only provides service to class 2 is that demand of class 2 during the lead time ($D_1(\tau) = 0$ for any $\tau > 0$) is less or equal to the reorder point r, i.e., $D_2(L) \leq r$. Therefore, the service level provided to the high priority class when the DC only provides service to class 2 is:

$$\alpha_2(r, C) = \mathbb{P}(D_2(L) \leq r). \tag{B.2}$$

Equation (5.2) is equal to equation (B.2) when $D_2(\tau) > 0$ for any $\tau > 0$, $D_1(\tau) = 0$ for any $\tau > 0$ and $C = 0$. In subsection 5.2.2 we show that when $D_1(\tau) = 0$ for any $\tau > 0$, then the optimal critical level is equal to zero, i.e., $C^* = 0$. Furthermore, in subsection 5.2.2 we show that in the absence of demand for class 2, $\alpha_2(r^*, C^*) > \overline{\alpha}_2$. Therefore, the equation (5.2) is a general expression for the service level provided to the low priority class. □

B.3 Proof of proposition 9

Proof. Let $C_2(r)$ be the maximum critical level, given a reorder point r, that ensures a service level $\overline{\alpha}_2$, i.e., $C_2(r) = \max\{C \mid \alpha_2(r, C) \geq \overline{\alpha}_2\}$. In chapter 3 we showed that $\alpha_2(r, C)$ is increasing in r (and decreasing in C), therefore, $C_2(r)$ is solution of $\alpha_2(r, C) = \overline{\alpha}_2$. Then, $C_2(r) = r - r_2^0 = r - \mu L - z_{\overline{\alpha}_2}\sigma\sqrt{L}$. In the same way we define $C_1(r)$ as the minimum critical level, given a reorder point r, that ensures a service level $\overline{\alpha}_1$, i.e., $C_1(r) = \min\{C \mid \alpha_1(r, C) \geq \overline{\alpha}_1\}$. Let r_i^0 be the minimum reorder point r such that the service level provided to the class i, given a critical level $C = 0$, is greater than or equal to his preset service level $\overline{\alpha}_i$, i.e., $r_i^0 = \min\{r \mid \alpha_i(r, 0) \geq \overline{\alpha}_i\}$, with $i = 1, 2$.

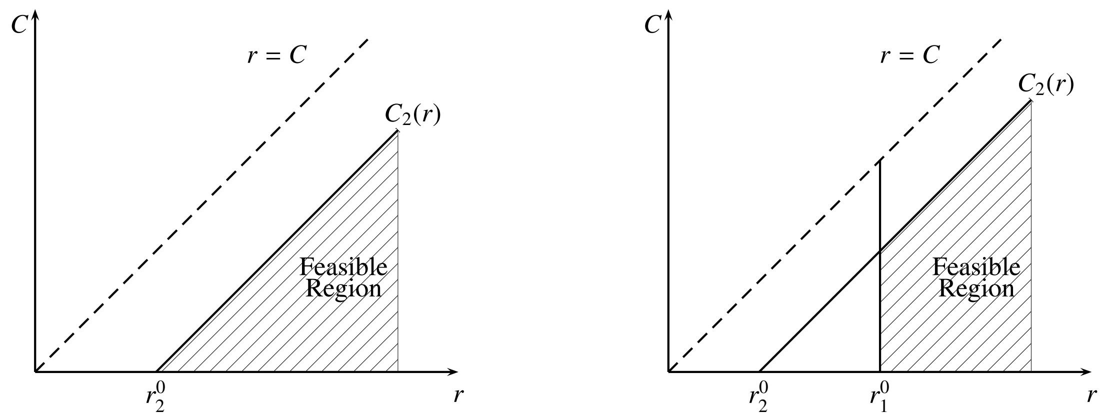
(a) If $\mu_1 = \sigma_1^2 = 0$, from equation (5.1) we have for any $r > C \geq 0$ that $\alpha_1(r, C) = \mathbb{P}(D_2(L) \leq r - C) + \mathbb{P}(\tau_{H,D_2}^{r-C} < L) = \mathbb{P}(D_2(L) \leq r - C) + 1 - \mathbb{P}(D_2(L) \leq r - C) = 1$. Therefore, $\alpha_1(r, C) = 1$, $\forall C \geq 0$ and $C_1(r) = 0$, $\forall r$. On the other hand, $C_2(r) = r - \mu_2 L - z_{\overline{\alpha}_2}\sigma_2\sqrt{L}$. Once $C_1(r)$ and $C_2(r)$ are defined, the feasible region of SLP problem where all (r, C) satisfies $\alpha_1(r, C) \geq \overline{\alpha}_1$, $\alpha_2(r, C) \geq \overline{\alpha}_2$ and $r > C \geq 0$, when $\mu_1 = \sigma_1^2 = 0$, is the intersection of the areas above $C_1(r)$, below $C_2(r)$ and strictly below $r = C$. The feasible region is shown in figure (B.1a).

From figure (B.1a) we conclude that the minimum reorder point that guarantees a service level $\overline{\alpha}_2$ provided to the low priority class is r_2^0. Therefore, $r^* = r_2^0 = \mu_2 L + z_{\overline{\alpha}_2}\sigma_2\sqrt{L}$ and $C^* = 0$.

(b) If $\mu_2 = \sigma_2^2 = 0$, it valid that $\alpha_1(r, C) = \alpha_1(r, 0)$, $\forall C \geq 0$. Therefore, $C_1(r) = 0$ for any $r \geq r_1^0$. On the other hand, $C_2(r) = r - r_2^0 = r - \mu_1 L - z_{\overline{\alpha}_2}\sigma_1\sqrt{L}$. The feasible region of SLP problem where all (r, C) satisfies $\alpha_1(r, C) \geq \overline{\alpha}_1$, $\alpha_2(r, C) \geq \overline{\alpha}_2$ and $r > C \geq 0$, when $\mu_2 = \sigma_2^2 = 0$, is the intersection of the areas above $C_1(r)$, below $C_2(r)$ and strictly below $r = C$. The feasible region is shown in figure (B.1b).

From figure (B.1b) we conclude that the minimum reorder point that guarantees a service level $\overline{\alpha}_1$ provided to the high priority class is r_1^0. Therefore, $r^* = r_1^0$ solution of $\alpha_1(r, 0) = \overline{\alpha}_1$ and for convenience $C^* = 0$. Note that $\alpha_2(r^*, 0) > \overline{\alpha}_2$ because $r_2^0 < r_1^0$.

□



(a) Feasible region of SLP problem when $\mu_1 = \sigma_1^2 = 0$ and $\mu_2, \sigma_2^2 > 0$

(b) Feasible region of SLP problem when $\mu_2 = \sigma_2^2 = 0$ and $\mu_1, \sigma_1^2 > 0$

Figure B.1: Feasible region of SLP problem under normally distributed demand and single class demand.

B.4 Numerical results for test sets

		Parameters			Solution approach														△cost
					First solution			Second solution			Critical level policy					Global Round-up			
Data Set	Vary	$\overline{\alpha}_2$	h_j	c_{ij}	$\sharp\, DC$	$FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2})$	Time(s)	$\sharp\, DC$	$FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})$	Time(s)	LB	FO^*	$Gap(\%)$	$\sharp\, DC$	Time(s)	$FO_{P1}(\overline{\alpha}_1)$	$\sharp\, DC$	Time(s)	(%)
49-nodes	Case base	0.75	0.25	0.01	12	8766.65	1.40	12	8766.65	1.78	8709.91	8766.65	0.65	12	3.18	8819.29	12	1.79	0.60
	$\overline{\alpha}_2$	**0.55**	0.25	0.01	12	8745.29	1.35	12	8745.29	1.77	8663.21	8745.29	0.94	12	3.11	8819.29	12	1.79	0.84
		0.65	0.25	0.01	12	8755.40	1.52	12	8755.40	1.79	8685.30	8755.40	0.80	12	3.32	8819.29	12	1.79	0.72
		0.85	0.25	0.01	12	8780.73	1.42	12	8780.73	1.77	8740.71	8780.73	0.46	12	3.19	8819.29	12	1.79	0.44
		0.95	0.25	0.01	12	8804.37	1.64	12	8804.37	1.76	8792.48	8804.37	0.14	12	3.41	8819.29	12	1.79	0.17
	h_j	0.75	**0.005**	0.01	12	6188.71	0.38	12	6188.71	0.31	6187.53	6188.71	0.02	12	0.69	6189.77	12	0.32	0.02
		0.75	**0.1**	0.01	12	7642.65	0.47	12	7642.65	0.50	7619.95	7642.65	0.30	12	0.97	7663.70	12	0.50	0.27
		0.75	**0.35**	0.01	11	9320.38	2.76	10	9318.37	3.04	9242.65	9318.37	0.81	10	5.80	9386.98	10	3.03	0.73
		0.75	**0.50**	0.01	10	9992.74	10.15	9	9995.60	16.86	9885.64	9992.74	1.07	10	27.02	10088.43	9	16.72	0.95
		0.75	**0.75**	0.01	9	10890.68	43.02	9	10890.68	67.19	10733.56	10890.68	1.44	9	110.22	11029.92	9	68.04	1.26
		0.75	**1.00**	0.01	7	11589.59	163.53	7	11589.59	141.39	11410.71	11589.59	1.54	7	304.92	11761.83	7	140.84	1.46
		0.75	**1.25**	0.01	7	12206.26	233.89	6	12197.85	145.95	11982.67	12197.85	1.76	6	379.85	12387.38	6	147.86	1.53
	c_{ij}	0.75	0.25	**0.005**	6	6558.11	6.30	6	6558.11	7.74	6517.31	6558.11	0.62	6	14.03	6597.14	6	7.76	0.59
		0.75	0.25	**0.015**	13	10476.10	0.81	12	10475.07	0.83	10415.98	10475.07	0.56	12	1.64	10527.71	12	0.82	0.50
		0.75	0.25	**0.02**	18	11552.20	1.16	18	11552.20	1.27	11487.46	11552.20	0.56	18	2.43	11620.44	18	1.30	0.59
	$\overline{\alpha}_2, h_j$	**0.55**	**0.5**	0.01	10	9952.95	9.61	9	9957.92	16.44	9798.06	9952.95	1.56	10	26.05	10088.43	9	16.72	1.34
		0.55	**1.00**	0.01	7	11519.14	238.56	7	11519.14	140.71	11260.81	11519.14	2.24	7	379.26	11761.83	7	139.37	2.06
		0.55	**1.25**	0.01	7	12118.20	450.61	6	12120.36	147.49	11795.28	12118.20	2.66	7	598.10	12387.38	6	147.16	2.17
		0.95	**0.5**	0.01	9	10062.13	14.57	9	10062.13	16.57	10040.00	10062.13	0.22	9	31.14	10088.43	9	16.71	0.26
		0.95	**1.00**	0.01	7	11713.94	201.50	7	11713.94	140.23	11675.76	11713.94	0.33	7	341.72	11761.83	7	139.37	0.41
		0.95	**1.25**	0.01	6	12334.61	226.86	6	12334.61	146.32	12292.57	12334.61	0.34	6	373.17	12387.38	6	147.41	0.43
	$\overline{\alpha}_2, c_{ij}$	**0.55**	0.25	**0.005**	6	6542.14	5.57	6	6542.14	7.75	6483.23	6542.14	0.90	6	13.32	6597.14	6	7.73	0.83
		0.55	0.25	**0.015**	12	10243.91	0.82	12	10243.91	0.71	10161.83	10243.91	0.80	12	1.53	10317.91	12	0.73	0.72
		0.55	0.25	**0.02**	19	11531.42	1.33	18	11524.42	1.34	11429.38	11524.42	0.82	18	2.67	11620.44	18	1.28	0.83
		0.95	0.25	**0.005**	6	6586.31	6.75	6	6586.31	7.69	6577.57	6586.31	0.13	6	14.44	6597.14	6	7.76	0.16
		0.95	0.25	**0.015**	12	10302.99	0.77	12	10302.99	0.71	10291.10	10302.99	0.12	12	1.48	10317.91	12	0.71	0.14
		0.95	0.25	**0.02**	18	11601.27	1.18	18	11601.27	1.26	11587.85	11601.27	0.12	18	2.44	11620.44	18	1.29	0.16
	h_j, c_{ij}	0.75	**1.25**	**0.005**	4	9023.39	1477.19	4	9023.39	3508.15	8858.55	9023.39	1.83	4	4985.34	9193.62	4	3537.09	1.85
		0.75	**1.25**	**0.015**	12	14271.58	21.546	11	14264.62	24.368	13987.86	14264.62	1.94	11	45.91	14521.85	11	24.973	1.77
		0.75	**1.25**	**0.02**	12	15770.25	2.16	12	15770.25	1.71	15486.53	15770.25	1.80	12	3.88	16033.43	12	1.796	1.64

Data Set	Vary	Parameters			First solution			Second solution			Critical level policy					Global Round-up			$\triangle cost$
		$\overline{\alpha}_2$	h_j	c_{ij}	$\sharp DC$	$FO_{P0}(X^*_{\overline{\alpha}_2}, Y^*_{\overline{\alpha}_2})$	Time(s)	$\sharp DC$	$FO_{P0}(X^*_{\overline{\alpha}_1}, Y^*_{\overline{\alpha}_1})$	Time(s)	LB	FO^*	$Gap(\%)$	$\sharp DC$	Time(s)	$FO_{P1}(\overline{\alpha}_1)$	$\sharp DC$	Time(s)	(%)
88-nodes	Case base	0.75	0.25	0.01	15	15831.84	14.80	15	15831.84	23.98	15750.57	15831.84	0.51	15	38.78	15914.42	15	24.45	0.52
	$\overline{\alpha}_2$	**0.55**	0.25	0.01	15	15798.22	20.80	15	15798.22	24.31	15680.62	15798.22	0.74	15	45.10	15914.42	15	24.45	0.73
		0.65	0.25	0.01	15	15814.13	15.64	15	15814.13	24.16	15713.71	15814.13	0.63	15	39.80	15914.42	15	24.45	0.63
		0.85	0.25	0.01	15	15853.98	16.64	15	15853.98	24.22	15796.70	15853.98	0.36	15	40.86	15914.42	15	24.45	0.38
		0.95	0.25	0.01	15	15891.18	18.49	15	15891.18	24.03	15874.25	15891.18	0.11	15	42.52	15914.42	15	24.45	0.15
	h_j	0.75	**0.005**	0.01	21	11549.66	1.49	21	11549.66	1.42	11547.71	11549.66	0.02	21	2.91	11551.56	21	1.40	0.02
		0.75	**0.05**	0.01	20	13105.64	2.24	20	13105.64	2.23	13087.70	13105.64	0.14	20	4.47	13123.96	20	2.24	0.14
		0.75	**0.1**	0.01	19	14052.27	3.71	19	14052.27	4.04	14016.88	14052.27	0.25	19	7.75	14087.74	19	3.97	0.25
		0.75	**0.40**	0.01	13	17015.56	47.11	11	17010.60	31.30	16890.26	17010.60	0.71	11	78.42	17126.94	11	31.345	0.68
		0.75	**0.5**	0.01	11	17623.92	38.89	11	17623.92	53.30	17477.97	17623.92	0.83	11	92.19	17769.34	11	54.37	0.82
		0.75	**0.75**	0.01	11	18951.74	117.66	11	18951.74	220.41	18732.82	18951.74	1.16	11	338.07	19169.86	11	220.11	1.14
		0.75	**1.00**	0.01	11	20094.59	771.74	10	20094.23	4414.24	19802.70	20094.23	1.45	10	5185.98	20376.72	10	4320.96	1.39
		0.75	**1.25**	0.01	10	21059.36	2608.21	10	21059.36	10812.15[a]	20717.83	21059.36	1.62	10	13420.36	21412.47	10	10809.80[a]	1.65
	c_{ij}	0.75	0.25	**0.005**	10	11719.49	72.59	10	11719.49	116.68	11651.18	11719.49	0.58	10	189.27	11790.11	10	118.12	0.60
		0.75	0.25	**0.015**	23	18449.29	8.43	22	18442.41	10.19	18346.98	18442.41	0.52	22	18.62	18538.83	22	10.16	0.52
		0.75	0.25	**0.02**	28	20274.77	2.49	28	20274.77	2.72	20164.00	20274.77	0.55	28	5.21	20382.88	28	2.67	0.53
	$\overline{\alpha}_2, h_j$	**0.55**	**0.05**	0.01	20	13098.21	2.12	20	13098.21	2.23	13072.22	13098.21	0.20	20	4.36	13123.96	20	2.21	0.20
		0.55	**0.4**	0.01	13	16964.89	27.29	11	16962.91	30.51	16783.70	16962.91	1.06	11	57.81	17126.94	11	31.08	0.96
		0.55	**1.00**	0.01	11	19975.36	693.47	10	19978.26	4419.15	19553.91	19975.36	2.11	11	5112.62	20376.72	10	4385.18	1.97
		0.55	**1.25**	0.01	10	20914.40	2908.11	10	20914.40	10810.67[a]	20421.25	20914.40	2.36	10	13718.78	21412.47	10	10810.03[a]	2.33
		0.95	**0.05**	0.01	20	13118.74	2.21	20	13118.74	2.24	13115.07	13118.74	0.03	20	4.45	13123.96	20	2.21	0.04
		0.95	**0.4**	0.01	11	17094.85	34.68	11	17094.85	31.34	17069.80	17094.85	0.15	11	66.02	17126.94	11	31.38	0.19
		0.95	**1.00**	0.01	10	20299.07	2073.77	10	20299.07	4225.10	20240.49	20299.07	0.29	10	6298.87	20376.72	10	4330.09	0.38
		0.95	**1.25**	0.01	10	21315.41	10809.19[a]	10	21315.41	10809.53[a]	21242.19	21315.41	0.34	10	21618.73	21412.47	10	10809.46[a]	0.45
	$\overline{\alpha}_2, c_{ij}$	**0.55**	0.25	**0.005**	10	11690.49	48.93	10	11690.49	115.82	11591.86	11690.49	0.84	10	164.75	11790.11	10	117.62	0.84
		0.55	0.25	**0.015**	23	18409.14	6.99	22	18403.29	10.33	18260.95	18403.29	0.77	22	17.31	18538.83	22	10.32	0.73
		0.55	0.25	**0.02**	28	20230.91	2.49	28	20230.91	2.66	20070.54	20230.91	0.79	28	5.15	20382.88	28	2.67	0.75
		0.95	0.25	**0.005**	10	11770.70	200.16	10	11770.70	117.83	11756.05	11770.70	0.12	10	317.99	11790.11	10	116.97	0.16
		0.95	0.25	**0.015**	22	18511.46	9.17	22	18511.46	10.29	18492.58	18511.46	0.10	22	19.46	18538.83	22	10.27	0.15
		0.95	0.25	**0.02**	28	20352.21	2.89	28	20352.21	2.76	20329.23	20352.21	0.11	28	5.66	20382.88	28	2.72	0.15
	h_j, c_{ij}	0.75	**1.0**	**0.005**	7	15100.30	10812.81[a]	7	15100.30	10808.06[a]	14875.43	15100.30	1.49	7	22110.88	15337.88	7	10816.931[a]	1.55
		0.75	**1.0**	**0.015**	17	23868.85	3949.97	16	23862.61	10520.11	23535.58	23862.61	1.37	16	14470.09	24189.39	16	10274.19	1.35
		0.75	**1.0**	**0.02**	22	26387.26	370.98	22	26387.26	311.39	26018.35	26387.26	1.40	22	682.37	26772.94	22	311.59	1.44

Solution approach: First solution, Second solution, Critical level policy.

[a] Suboptimal solution obtained for 3 h limit.

Table B.2: Results for test set

C | Appendix Chapter 6

C.1 Pseudocodes

Algorithm 3 Pseudocode UB $\mathbf{CQ_{LRU}}$

for $i \in N_1$ **do**
Find j installed, serving class 1 and with smallest $\widehat{d_{ij}}$
Set $j_{min1} = j$
Set $y_{i,j_{min1}} = 1$
end for
for $i \in N_2$ **do**
Find j installed, serving class 2 and with smallest $\widehat{d_{ij}}$
Set $j_{min2} = j$
Set $y_{i,j_{min2}} = 1$
end for
Evaluate $V(\mathbf{CQ_{LRU}})$
Obtain $UB = V(\mathbf{CQ_{LRU}})$

Algorithm 4 Lagrangian Relaxation Algorithm

Set $k = 0$, $LB^* = -\infty$, $UB^* = \infty$
while $|UB^* - LB^*| > \epsilon$ **do**
Obtain $LB^{(k)}$ using (6.34)
if $LB^{(k)} \geq LB^*$ **then**
Set $LB^* = LB^{(k)}$
end if
Obtain $UB^{(k)}$ using Algorithm 3
if $UB^{(k)} \leq UB^*$ **then**
Set $UB^* = UB^{(k)}$
end if
Update $\lambda_i^{(k)}$
Set $k = k + 1$
end while

C.2 Numerical results for test sets

		Parameters			Global Round Up			Local Round Up				Single Class Allocation			Separate stock			Critical level policy			
Data Set	Vary	β_2	h_j	c_{ij}	♯ DC	Z_{GRU}	Time(s)	♯ DC	Z_{LRU}	Time(s)	(%)gap	♯ DC	Z_{SCA}	Time(s)	♯ DC	Z_{SSP}	Time	♯ DC	LB	UB	Time(s)
49-nodes	Case base	0.750	0.250	0.010	12	8819.29	1.79	12	8813.02	372.40	0.000	16	9982.83	13.58	12	8795.97	1.53	12	8709.91	8766.65	3.18
	β_2	0.550	0.250	0.010	12	8819.29	1.79	12	8810.39	514.66	0.000	16	9957.41	13.49	12	8768.18	1.70	12	8663.21	8745.29	3.11
		0.650	0.250	0.010	12	8819.29	1.80	12	8811.63	415.14	0.000	16	9969.44	15.64	12	8781.33	1.74	12	8685.30	8755.40	3.32
		0.850	0.250	0.010	12	8819.29	1.78	12	8819.18	365.12	0.071	16	9999.59	13.38	12	8814.30	1.68	12	8740.71	8780.73	3.19
		0.950	0.250	0.010	12	8819.29	1.79	12	8817.67	507.46	0.005	16	10027.77	12.59	12	8845.11	1.47	12	8792.48	8804.37	3.41
	h_j	0.750	0.005	0.010	12	6189.77	0.32	13	6189.84	556.83	0.055	19	6910.10	0.47	12	6189.30	0.34	12	6187.53	6188.71	0.69
		0.750	0.100	0.010	12	7663.70	0.52	12	7661.13	311.66	0.008	16	8675.86	2.05	12	7654.38	0.50	12	7619.95	7642.65	0.97
		0.750	0.350	0.010	10	9386.98	3.07	11	9390.62	495.16	0.099	14	10640.84	400.89	11	9359.54	3.80	10	9242.65	9318.37	5.80
		0.750	0.500	0.010	9	10088.43	16.88	9	10090.01	725.18	0.016	12	11403.47	10800.388[a]	10	10056.96	17.99	10	9885.64	9992.74	27.02
		0.750	0.750	0.010	9	11029.92	66.71	7	11033.80	1219.67	0.099	10	12348.67	5183.30	9	10985.40	8514.41	9	10733.56	10890.68	110.22
		0.750	1.000	0.010	7	11761.83	137.55	7	11765.52	660.94	0.034	10	13170.57	11008.834[a]	7	11714.27	200.96	7	11410.71	11589.59	304.92
		0.750	1.250	0.010	6	12387.38	149.51	6	12389.35	759.92	0.016	8	13845.32	11030.123[a]	6	12340.35	245.95	6	11982.67	12197.85	379.85
	C_{ij}	0.750	0.250	0.005	6	6597.14	7.74	6	6597.59	669.90	0.007	8	7509.06	138.05	6	6585.74	6.47	6	6517.31	6558.11	14.03
		0.750	0.250	0.015	12	10317.91	0.72	12	10311.73	321.00	0.000	19	11395.24	3.10	12	10294.60	0.70	12	10208.53	10265.27	1.49
		0.750	0.250	0.020	18	11620.44	1.32	18	11592.26	423.93	0.034	22	12416.53	1.48	18	11577.71	1.33	18	11487.46	11552.20	2.43
	β_2, h_j	0.550	0.500	0.010	9	10088.43	16.38	9	10090.01	580.74	0.020	11	11355.03	10800.309[a]	10	10000.47	17.16	10	9798.06	9952.95	26.05
		0.550	1.000	0.010	7	11761.83	135.31	7	11765.52	491.11	0.093	10	13077.69	11080.337[a]	7	11612.39	4437.61	7	11260.81	11519.14	379.26
		0.550	1.250	0.010	6	12387.38	146.00	6	12389.35	444.33	0.016	8	13736.82	11371.077[a]	6	12229.14	476.42	7	11795.28	12118.20	598.10
		0.950	0.500	0.010	9	10088.43	16.26	9	10090.01	621.93	0.016	11	11474.06	1422.89	9	10153.79	16.63	9	10040.00	10062.13	31.14
		0.950	1.000	0.010	7	11761.83	135.54	7	11765.52	515.56	0.056	9	13327.20	10952.904[a]	7	11894.40	240.91	7	11675.76	11713.94	341.72
		0.950	1.250	0.010	6	12387.38	162.76	6	12389.35	534.35	0.016	8	14037.15	10917.027[a]	6	12536.97	296.97	6	12292.57	12334.61	373.17
	β_2, C_{ij}	0.550	0.250	0.005	6	6597.14	7.76	6	6597.59	702.99	0.007	8	7487.36	109.10	6	6563.50	8.01	6	6483.23	6542.14	13.32
		0.550	0.250	0.015	12	10317.91	0.70	12	10309.09	318.91	0.000	19	11364.97	2.90	12	10266.80	0.74	12	10161.83	10243.91	1.53
		0.550	0.250	0.020	18	11620.44	1.31	18	11578.00	362.78	0.031	22	12384.69	1.41	18	11544.44	1.42	18	11429.38	11524.42	2.67
		0.950	0.250	0.005	6	6597.14	7.71	6	6597.59	620.40	0.007	8	7547.43	134.87	5	6624.34	8.23	6	6577.57	6586.31	14.44
		0.950	0.250	0.015	12	10317.91	0.71	12	10316.38	248.14	0.121	19	11448.77	3.45	12	10343.73	0.70	12	10291.10	10302.99	1.48
		0.950	0.250	0.020	18	11620.44	1.32	18	11617.47	383.12	0.092	22	12472.81	1.30	18	11636.53	1.43	18	11587.85	11601.27	2.44
	h_j, C_{ij}	0.750	1.250	0.005	4	9193.62	3644.97	4	9193.59	537.04	0.161	5	10412.83	10800.318[a]	4	9150.76	2234.08	4	8858.55	9023.39	4985.34
		0.750	1.250	0.015	11	14521.85	24.28	12	14503.93	456.41	0.000	15	16226.68	4642.50	12	14418.18	42.13	11	13987.86	14264.62	45.91
		0.750	1.250	0.020	12	16033.43	1.69	12	16002.63	324.10	0.000	18	17650.49	67.13	12	15916.85	1.99	12	15486.53	15770.25	3.88

[a] Suboptimal solution obtained for 3 h limit.

		Parameters			Global Round Up			Local Round Up				Single Class Allocation			Separate stock			Critical level policy			
Data Set	Vary	β_2	h_j	c_{ij}	$\sharp DC$	Z_{GRU}	Time(s)	$\sharp DC$	Z_{LRU}	Time(s)	(%)gap	$\sharp DC$	Z_{SCA}	Time(s)	$\sharp DC$	Z_{SSP}	Time	$\sharp DC$	LB	UB	Time(s)
88-nodes	Case base	0.750	0.250	0.010	15	15914.42	23.88	15	15914.42	893.433	0.089	19	18605.93	813.43	15	15893.80	20.64	15	15750.57	15831.84	23.98
	β_2	0.550	0.250	0.010	15	15914.42	24.51	15	15914.42	1050.204	0.093	19	18566.99	952.77	15	15848.04	18.55	15	15680.62	15798.22	24.31
		0.650	0.250	0.010	15	15914.42	23.84	15	15914.42	1087.851	0.093	19	18585.41	960.53	15	15869.69	20.85	15	15713.71	15814.13	24.16
		0.850	0.250	0.010	15	15914.42	24.61	15	15914.42	1097.435	0.091	19	18631.61	832.88	15	15923.98	20.36	15	15796.70	15853.98	24.22
		0.950	0.250	0.010	15	15914.42	23.78	15	15914.42	998.654	0.089	19	18674.78	2136.72	15	15974.71	24.31	15	15874.25	15891.18	24.03
	h_j	0.750	0.005	0.010	21	11551.56	1.43	20	11565.00	750.779	0.126	26	14036.30	2.29	21	11550.92	1.44	21	11547.71	11549.66	1.42
		0.750	0.100	0.010	19	14087.74	4.01	19	14089.97	1065.931	0.029	21	16673.70	28.12	19	14077.51	3.93	19	14016.88	14052.27	4.04
		0.750	0.350	0.010	13	16769.94	42.34	13	16769.94	1637.537	0.000	18	19559.21	8555.95	13	16744.00	53.30	13	16551.55	16661.19	84.42
		0.750	0.500	0.010	11	17769.34	55.24	11	17769.34	1086.833	0.082	17	20744.98	10962.41[a]	11	17736.74	45.34	11	17477.97	17623.92	53.30
		0.750	0.750	0.010	11	19169.86	217.92	11	19169.86	1353.314	0.067	16	22371.99	10813.32[a]	11	19120.97	176.01	11	18732.82	18951.74	220.41
		0.750	1.000	0.010	10	20376.72	4953.21	10	20376.72	1395.378	0.016	15	23733.96	10964.12[a]	10	20318.57	3014.58	11	19802.70	20094.23	5185.98
		0.750	1.250	0.010	10	21412.47	10812.62[a]	10	21412.47	908.486	0.005	14	24879.66	10909.89[a]	10	21339.78	9398.08	10	20717.83	21059.36	10812.15[a]
	C_{ij}	0.750	0.250	0.005	10	11790.11	114.49	10	11781.57	680.967	0.005	14	13962.29	10801.8[a]	9	11783.86	10800.40[a]	10	11651.18	11719.49	116.68
		0.750	0.250	0.015	22	18538.83	10.33	22	18534.04	740.848	0.075	28	21883.62	221.32	22	18504.10	9.27	23	18346.98	18442.41	10.19
		0.750	0.250	0.020	28	20382.88	2.71	28	20361.57	645.032	0.036	40	23792.41	34.72	28	20332.13	3.06	28	20164.00	20274.77	2.72
	β_2, h_j	0.550	0.500	0.010	11	17769.34	54.21	11	17769.34	911.339	0.073	16	20699.34	10802.864[a]	11	17653.07	36.23	11	17353.58	17564.30	89.62
		0.550	1.000	0.010	10	20376.72	4490.93	10	20376.72	1454.485	0.059	15	23614.70	10987.232[a]	11	20152.89	1379.38	11	19553.91	19975.36	4419.15
		0.550	1.250	0.010	10	21412.47	10810.86[a]	10	21412.47	1239.571	0.070	15	24728.21	10890.89[a]	10	21135.88	5048.98	10	20421.25	20914.40	10810.67[a]
		0.950	0.500	0.010	11	17769.34	54.49	11	17769.34	1482.631	0.073	17	20872.78	11070.09[a]	11	17884.68	46.39	11	17697.91	17729.23	95.78
		0.950	1.000	0.010	10	20376.72	4646.05	10	20376.72	1689.813	0.046	15	24020.24	10863.60[a]	10	20606.97	10800.75[a]	10	20240.49	20299.07	4225.10
		0.950	1.250	0.010	10	21412.47	10811.81[a]	10	21412.47	1080.587	0.001	13	25177.33	10956.82[a]	10	21699.98	10809.79[a]	10	21242.19	21315.41	10809.53[a]
	β_2, C_{ij}	0.550	0.250	0.005	10	11790.11	119.83	10	11777.12	2560.345	0.069	13	13894.60	10802.91[a]	10	11734.79	105.58	10	11591.86	11690.49	115.82
		0.550	0.250	0.015	22	18538.83	10.37	22	18528.95	1308.505	0.078	28	21840.58	157.11	22	18451.71	10.31	23	18260.95	18403.29	10.33
		0.550	0.250	0.020	28	20382.88	2.68	29	20354.71	2584.432	0.030	40	23738.52	23.27	28	20275.31	2.71	28	20070.54	20230.91	2.66
		0.950	0.250	0.005	10	11790.11	130.20	10	11794.77	1809.364	0.036	13	13945.97	10848.296[a]	9	11860.61	10825.35[a]	10	11756.05	11770.70	117.83
		0.950	0.250	0.015	22	18538.83	10.34	22	18543.05	2409.453	0.050	28	21959.71	189.12	22	18596.73	10.25	22	18492.58	18511.46	10.29
		0.950	0.250	0.020	28	20382.88	2.71	29	20385.47	1012.548	0.072	40	23887.69	19.47	28	20432.57	3.51	28	20329.23	20352.21	2.76
	h_j, C_{ij}	0.750	1.250	0.005	6	16117.16	10824.96[a]	7	15887.24	2070.478	0.068	10	18937.21	10801.701[a]	6	16049.31	10827.02[a]	6	15653.97	15837.63	10812.99[a]
		0.750	1.250	0.015	16	25459.96	10800.54[a]	16	25429.84	2146.011	0.050	19	29244.21	10809.53[a]	17	25390.37	10822.34[a]	16	24644.08	25051.48	10802.19[a]
		0.750	1.250	0.020	20	28257.18	2740.27	21	28228.63	2036.999	0.097	27	32563.90	10810.67[a]	21	28088.44	1487.33	20	27327.69	27802.25	4697.15

[a] Suboptimal solution obtained for 3 h limit.

Table C.1: Results for test set

D | Appendix Chapter 7

D.1 Data for illustrative example

Candidate DCs				Customers					
	Location		f_j			Location		Demand	
DC	X	Y	$(US\$/day)$	Customer	Class	X	Y	μ_i (kg/day)	CV_i
1	-9.47	21.02	191	1	1	-7.37	10.71	4647.33	0.53
2	-6.22	15.53	200	2	1	-18.23	0.11	1616.67	0.42
3	-6.67	15.92	195	3	1	-13.31	-12.01	1275.33	0.85
4	-7.98	16.39	206	4	2	-9.38	10.96	504.67	0.63
5	-8.60	16.71	206	5	2	-11.15	-3.24	430.00	0.59
6	-7.59	16.18	202	6	1	-5.27	1.55	882.00	0.93
7	-7.56	15.23	211	7	2	-6.13	4.23	248.00	0.36
8	-7.10	14.90	206	8	1	2.19	1.11	4880.00	0.78
9	-6.45	14.93	211	9	2	4.13	-0.22	386.67	0.78
10	-6.89	14.03	206	10	1	-8.16	13.68	1256.67	0.44
11	-5.97	13.62	244	11	1	-13.15	5.45	1476.67	0.61
12	-4.91	11.62	237	12	2	-2.48	6.04	176.00	0.17
13	-3.56	12.01	233	13	1	-8.32	10.19	700.67	0.15
14	-0.89	10.31	233	14	2	7.18	4.45	230.00	0.58
15	-4.49	9.56	255	15	1	-6.40	5.12	872.67	0.80
16	-4.35	9.29	255	16	2	-8.81	-9.94	196.00	0.94
17	-3.64	5.38	233	17	2	-7.05	0.22	388.00	0.22
18	-6.87	10.81	211	18	2	1.18	-9.12	176.00	0.61
19	-9.55	10.16	222	19	2	-11.24	-38.33	456.67	0.52
20	-10.27	10.08	211	20	2	-6.86	-12.23	114.00	0.11
21	-10.64	10.15	206	21	2	-2.61	-6.45	156.00	0.40
22	-10.65	9.79	217	22	2	-7.05	-18.35	250.67	0.25
23	-11.08	9.89	211	23	2	7.65	0.00	144.67	0.81
24	-11.87	9.55	203	24	2	-7.51	-18.12	321.33	0.38
25	-12.34	8.76	208	25	2	-10.55	-33.96	194.67	0.58
26	-22.13	2.04	200	26	2	4.04	-18.35	174.67	0.25
27	-12.88	2.72	222	27	2	-23.59	-19.68	168.00	0.64
28	-13.07	2.92	222	28	2	-2.71	-2.78	198.00	0.34
29	-13.10	2.53	211	29	2	-28.49	-26.02	150.67	0.69
30	-13.09	2.14	195	30	2	-12.41	-7.45	111.33	0.72
31	-12.32	2.31	211	31	2	9.77	7.78	100.00	0.77
32	-9.31	1.59	228	32	2	6.07	-6.12	88.67	0.51
33	-8.72	1.51	228	33	2	-8.59	-10.30	186.00	0.18
34	-6.93	-6.22	228	34	2	-4.37	7.78	154.00	0.31
35	-5.07	-3.42	217	35	2	-26.55	-25.24	264.67	0.92
36	-6.47	-8.86	222	36	2	2.56	-9.90	233.33	0.24
37	-6.98	-9.36	208	37	2	4.33	3.14	222.00	0.84
38	-6.49	-9.47	208	38	2	4.23	-20.35	109.33	0.58

Table D.1: Data for illustrative example