Contributions to the multi-echelon inventory optimisation problem using the guaranteed-service model approach

Ayse Sena Eruguz

ÉCOLE CENTRALE DES ARTS ET MANUFACTURES « ÉCOLE CENTRALE PARIS »

THÈSE

présentée par

Ayse Sena ERUGUZ

pour l'obtention du

GRADE DE DOCTEUR

Spécialité : **Génie Industriel**

Laboratoire d'accueil : **Laboratoire Génie Industriel**

CONTRIBUTIONS TO THE MULTI-ECHELON INVENTORY OPTIMISATION PROBLEM USING THE GUARANTEED-SERVICE MODEL APPROACH

Soutenue le : 13 février 2014

devant un jury composé de :

M. Vincent GIARD, Université Paris-Dauphine	**Président**
M. Haoxun CHEN, Université de Technologie Troyes	**Rapporteur**
M. Jean-Claude HENNET, CNRS-LSIS Marseille	**Rapporteur**
M. Stefan MINNER, Technische Universität München	**Examinateur**
M. Yves DALLERY, Ecole Centrale Paris	**Encadrant de thèse**
M. Zied JEMAI, Ecole Centrale Paris	**Encadrant de thèse**
Mme. Evren SAHIN, Ecole Centrale Paris	**Encadrante de thèse**

2014ECAP0020

Acknowledgments

First and foremost, I would like to thank my supervisors Zied JEMAI, Evren SAHIN and Yves DALLERY for their guidance, help and motivation. This thesis would have been impossible without their support and encouragement.

I also would like to thank the jury members Profs. Haoxun CHEN, Vincent GIARD, Jean-Claude HENNET and Stefan MINNER for their valuable comments and for initiating interesting and useful discussions at my defense.

I am also thankful to all my friends and colleagues at Ecole Centrale Paris who gave me their support and created an amazing work environment during these three years. I can deem myself extremely lucky for having been part of such a great team.

Finally, I would like to thank my parents and my brother for constantly offering me unconditional support, encouragement and love. We are spread out all over the planet, but they are always close to me. This work is dedicated to them.

Abstract

Many real-world supply chains can be characterised as large and complex multi-echelon systems since they consist of several stages incorporating assembly and distribution processes. A challenge facing such systems is the efficient management of inventory when demand is uncertain, operating costs and customer service requirements are high. This requires specifying the inventory levels at different stages that minimise the total cost and meet target customer service levels. In order to address this problem, researchers proposed the Stochastic-Service Model and the Guaranteed-Service Model (GSM) approaches. These two approaches differ in terms of assumptions with regard to how to address demand variations and service times.

This thesis develops several contributions to the GSM based multi-echelon inventory optimisation problem. First of all, we conduct a comprehensive literature review which gives a synthesis of the various GSM work developed so far. Then, we study the impact of some specific assumptions of the GSM such as bounded demand, guaranteed-service times and common review periods. Our numerical analysis shows that the bounded demand assumption may cause a deviation on customer service levels while the guaranteed-service times and common review periods assumptions may result in an increase on the total cost. In real-world supply chains the impact of these assumptions might be significant. Based on the findings presented while investigating the impact of the common review periods assumption, we develop an extension of the GSM that enables to simultaneously optimise the review periods (reorder intervals) and safety stock levels (order-up-to levels) in general acyclic multi-echelon systems. We formulate this problem as a nonlinear integer programming model. Then, we propose a sequential optimisation procedure that enables to obtain near optimal solutions with reasonable computational time. Finally, we focus on the issue of customer service level deviation in the GSM and propose two approaches in order to mitigate this deviation. The numerical study shows that the first approach outperforms the second one in terms of computational time while the second approach provides more accurate solutions in terms of cost. We also present some related issues in decentralised supply chain settings.

Keywords: Inventory control; Multi-echelon system; Guaranteed-service model; Optimisation; Supply chain

Résumé

De nombreuses chaînes logistiques peuvent être caractérisées comme de larges systèmes multi-échelons composés de plusieurs étages qui intègrent des activités d'assemblage et de distribution. L'un des enjeux majeurs associé au management de ces systèmes multi-échelons est la gestion efficace de stocks surtout dans des environnements où la demande est incertaine, les coûts de stocks sont importants et les exigences en termes de niveau de service client sont élevées. Cela nécessite en particulier de spécifier les niveaux de stocks aux différents étages afin de minimiser le coût total du système global et de satisfaire les niveaux cibles de service client. Pour faire face à ce problème, deux approches existent dans la littérature; il s'agit du Modèle de Service Stochastique (SSM) et le Modèle de Service Garanti (GSM). Ces deux approches diffèrent en termes d'hypothèses utilisées concernant la façon de gérer les variations de la demande et les temps de service.

Cette thèse amène plusieurs contributions au problème d'optimisation de stocks multi-échelons basé sur le GSM. Tout d'abord, nous menons une revue de la littérature internationale qui donne une synthèse des différents travaux réalisés à ce jour. Ensuite, nous étudions l'impact de certaines hypothèses spécifiques du GSM comme la demande bornée, les temps de service garanti et les périodes d'approvisionnement communes. Notre analyse numérique montre que l'hypothèse de demande bornée peut causer une déviation sur les niveaux de service client tandis que les hypothèses de temps de service garanti et de périodes d'approvisionnement communes peuvent entraîner une augmentation du coût total. En pratique, l'impact de ces hypothèses peut être important. En se basant sur les résultats présentés lors de l'analyse de l'hypothèse des périodes d'approvisionnement communes, nous développons une extension du GSM qui permet d'optimiser simultanément les périodes d'approvisionnement (les intervalles de réapprovisionnement) et les niveaux de stocks de sécurité (les niveaux de recomplétement) dans les systèmes multi-échelons acycliques généraux. Nous formulons ce problème comme un modèle de programmation non-linaire en nombres entiers. Ensuite, nous proposons une procédure d'optimisation séquentielle qui permet d'obtenir des solutions proches de l'optimal avec un temps de calcul raisonnable. Enfin, nous nous concentrons sur le problème de déviation de niveau de service client dans le GSM et nous proposons deux approches afin d'atténuer cette déviation. L'étude numérique montre que la première approche est plus performante que la deuxième en termes de temps de calcul tandis que la deuxième approche offre des meilleures solutions en termes de coût. Nous présentons également des problèmes similaires dans les chaînes logistiques décentralisées.

Mots-clés: Gestion de stocks; Systèmes multi-échelons; Modèle de service garanti; Optimisation; Chaîne logistique

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACSC	Annual Cycle Stock Cost
AFOC	Annual Fixed Ordering Cost
Appr.	Approximate
Ass.	Assembly
ASSC	Annual Safety Stock Cost
Avr.	Average
CoC	Clusters of Commonality
CSL	Cycle-Service-Level
DD	Default Direct
Distr.	Distribution
e.g.	for example
FCFS	First-Come-First-Served
FR	Fill-Rate
GO	Global Optimal
GSM	Guaranteed-Service Model
i.e.	that is
i.i.d.	identically and independently distributed
ID	Improved Direct
Max.	Maximum
Min.	Minimum
NLIP	Non Linear Integer Programming
PO2	Power-of-Two
Repl.	Replenishment
SO	Sequential Optimal
SOP	Sequential Optimisation Procedure
SSM	Stochastic-Service Model
vs.	versus

GENERAL INTRODUCTION

Motivations

A supply chain is a system that integrates several processes in order to convert the raw materials replenished from external suppliers into final products to be delivered to external customers. Supply chains may consist of several stages where each stage is associated with a process such as the procurement of a raw material, the production of a component, the manufacture of a subassembly, the assembly of a final product, its transportation from a central distribution centre to a regional warehouse or from a regional warehouse to a store (Graves and Willems, 2000). Indeed, many real-world supply chains can be characterised as large and complex multi-echelon systems since they may consist of thousands of stages incorporating both assembly and distribution processes. Several examples of such multi-echelon systems are illustrated by Willems (2008) for industries such as computer hard-ware, semiconductor, industrial chemicals, consumer goods and aircraft engine. A challenge facing these multi-echelon systems is the efficient management of inventory when demand is uncertain, operating costs are important and customer service requirements are high. This requires specifying the inventory decisions at different stages that minimise the total cost of the whole multi-echelon system and meet target customer service levels. In this thesis, we mainly focus on inventory decisions related to safety stock optimisation in multi-echelon supply chain systems.

Safety stock is introduced as a lever to cover uncertainties in inventory systems. In a multi-echelon system, the level of safety stock to be held at each stage must be suitably optimised in order to reach target customer service levels at the lowest cost. Indeed, the level of the local safety stock associated with each stage can be independently determined by using single-echelon inventory models which have widely been studied to date (see, e.g., Silver et al., 1998; Zipkin, 2000). However, such an approach would consider only the parameters associated with the relevant stage (e.g., the local inventory holding cost and processing lead time, target customer service levels, demand from the downstream stages, replenishment times from the upstream stages etc.). It would therefore lead to redundant safety stocks

because of the non-consideration of the interdependencies of cost and service level performances of connected stages. The multi-echelon safety stock optimisation approach aims at optimising safety stocks through a holistic view of the supply chain considering all stages in the supply chain simultaneously, from the external supplier to the external customer. Although the multi-echelon approach imposes significant computational challenges, it provides better results in comparison with the single-echelon approach in terms of cost and customer service level. Figure I illustrates the single- and multi-echelon approaches[1].

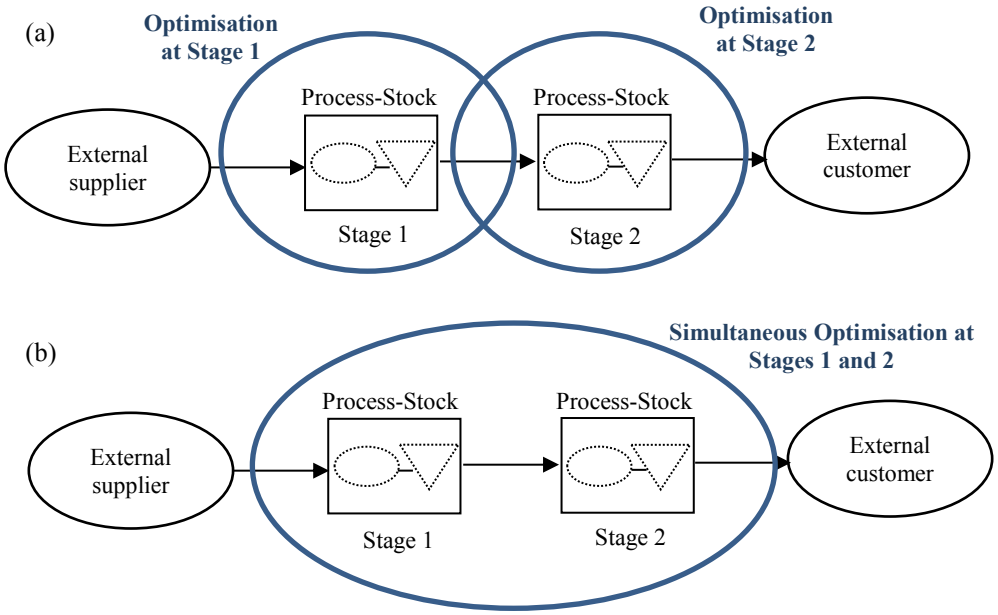

Figure I: Single-echelon approach (a), Multi-echelon approach (b)

According to Simchi-Levi and Zhao (2012) three reasons have contributed to the benefits obtained by a multi-echelon inventory optimisation approach: (1) the availability of data concerning demand and lead times, (2) motivations in industry to use scientific methods for inventory management, (3) recent developments in modelling and algorithms for the control of general multi-echelon structures. The benchmark report of Aberdeen Group (2007) also confirms the motivation of companies to use multi-echelon inventory optimisation tools. According to this report, in 2007, the 210 respondent companies ranked the inventory optimisation top on the list of investment areas where the multi-echelon inventory

[1] This figure is inspired from Klosterhalfen (2010)

optimisation topic was the top priority. In 2012, the estimated benefit from the implementation of multi-echelon inventory optimisation tools is presented as 3.1% service level improvement and 15% decrease in cash-to-cash cycle (Aberdeen Group, 2012).

The use of the multi-echelon inventory optimisation approach in order to allocate safety stocks under final customer demand uncertainty is widely studied in the literature. In order to deal with this problem, researchers proposed the Stochastic-Service Model (SSM) and the Guaranteed-Service Model (GSM) approaches that are introduced by Clark and Scarf (1960) and Simpson (1958), respectively. These two approaches differ in terms of assumptions made with regard to how to address demand variations and service times. Assumptions made in the GSM approach enable the consideration of real-world supply chains. However, the SSM approach mostly focuses on serial, assembly or two-echelon distribution systems and its deployment in industry is relatively limited. Indeed, the GSM approach has gained interest in recent years. The academic extensions of the GSM have rendered this approach more realistic. In parallel, efficient solution techniques have enabled to deal with the current large and complex multi-echelon structures. Hence, the GSM approach has enabled to realise important benefits in practice. For instance, Billington et al. (2004) have showed that savings realised by using the GSM approach for Hewlett-Packard's Digital Camera and Inkjet Supplies business exceeded $130 million. Farasyn et al. (2011) have reported that the GSM approach based multi-echelon models produced 7% of average inventory reduction at Procter & Gamble's business units. Wieland et al. (2012) have described a multi-echelon inventory optimisation project at Intel and indicated that after its implementation, inventory levels are reduced more than 11% providing average service levels exceeding 90%. In this thesis, we have a special focus on the GSM approach since both academicians and practitioners have recognised the practical, computational and economic advantages of this approach.

Contributions

This thesis deals with the multi-echelon inventory optimisation problem by using the GSM approach. We make several contributions in this direction.

Our first contribution is to provide a comprehensive literature review of the GSM approach. Indeed, in comparison with the SSM approach which is widely studied in the literature, the research on GSM approach has gained interest in the last decade. To the best of our knowledge, we did not identify a literature review which gives a synthesis of the various works developed so far. We present a comprehensive literature review by classifying the relevant papers along three axes: modelling assumptions considered, solution techniques developed and results obtained by industrial applications. Our literature review allows us identifying some gaps in the GSM literature and leads us to the research questions considered in this thesis. We present this literature review in Chapter 2. A preliminary version of this work is published in the proceedings of the *14th IFAC Symposium on Information Control Problems in Manufacturing*, *INCOM'12* (Eruguz et al., 2012).

From our literature review, we notice that the impacts of some specific assumptions of the GSM are not elaborately studied. In particular, a comprehensive analysis is required to analyse the cost and service level impact of assumptions regarding demand bounds, guaranteed-service times and common review periods. Hence, our second contribution in this thesis is to provide such an analysis by investigating the impact of each assumption separately. To do this, we mostly focus on serial and assembly systems. For the numerical analysis of each assumption, we consider the same test problems associated with a five-stage serial system presented previously by Schoenmeyr (2008) and a real-world assembly system examined by Graves and Willems (2003).

Concerning the analysis conducted in order to quantify the impact of the GSM assumptions, the first assumption examined states that demand is bounded at each stage of the supply chain. In practice, the demand bound at a stage represents the maximum amount of demand that can be satisfied from the stock of this stage during a certain coverage time. In the literature, demand bounds are usually specified using a safety factor that relates to a target Cycle-Service-Level (CSL). Our analysis shows that the effectively observed CSL at a stage that faces the external customer demand would usually be less than the target one. Under different target service levels, the relative service level deviation is 25% on average for the considered real-world system.

The second assumption examined implies that each stage quotes a guaranteed-service time to its customers and provides 100% service for these service times. Hence, backorders are not allowed between customer-supplier stages. Indeed, the total safety stock cost obtained under this assumption may be significantly higher than a solution obtained by solving a model without this assumption. For the real-world system considered in our numerical study, the guaranteed-service time assumption causes 42.2% safety stock cost increase.

The third assumption considered concerns the inventory control policy of the GSM. The original GSM assumes that each stage operates with a periodic-review, order-up-to policy with a common review period for all stages. In practice, review periods can differ from stage to stage considering the economies of scale and/or the availability of resources. We show how to incorporate stage-dependent nested review periods into the GSM. This requires developing an appropriate expression for the demand bounds of stages. The numerical analysis shows that the original GSM may represent a significant cost increase for long review periods. Under different review period profiles considered for the numerical study, the safety stock cost increase is up to 19.1%.

We present the results relative to the impact of the bounded demand, guaranteed-service times and common review periods assumptions in Chapter 3. The first part of this study (bounded demand assumption) is published in the proceedings of the *5th International Conference on Modeling, Simulation and Applied Optimization, ICMSAO'13* (Eruguz et al., 2013c). A preliminary version of this study is presented at the *11th ISIR Summer School on Research Trends in Inventory Management and Modeling* (Eruguz et al., 2013d). The consequences associated with the GSM assumptions demonstrate that there is a need for further model developments regarding the GSM approach.

In the GSM literature, existing models consider the review periods of stages as given input parameters of the multi-echelon safety stock optimisation problem. Our third contribution in this thesis is to provide and extension of the GSM by incorporating fixed ordering costs into the model in order to optimise the safety stock levels (order-up-to levels) and review periods (reorder intervals) simultaneously. To do this, we focus on nested Power-of-Two (PO2) reorder intervals due to their significant practical and computational

advantages. Under a nested PO2 policy, reorder intervals are power-of-two multiples and the reorder interval of a stage cannot be greater than the reorder intervals of its supplier stages. In order to compute the demand bound functions under this setting, we are inspired from the expression that we have proposed while analysing the common review periods assumption of the GSM. Under the existence of these demand bounds, we first propose a deterministic Non Linear Integer Programming (NLIP) model that determines nested PO2 reorder intervals and order-up-to levels in general acyclic multi-echelon systems. Second, by defining reasonable bounds for the decision variables of the NLIP model, we propose an improved direct approach that reduces the computational time in obtaining global optimal solutions while solving the NLIP model. Third, we propose a Sequential Optimisation Procedure (SOP) to obtain near optimal solutions with reasonable computational time. The numerical study demonstrates that for a general acyclic multi-echelon system with randomly generated parameters, the SOP is able to obtain near-optimal solutions of about 0.46% optimality gap on average in a few seconds. We present this contribution in Chapter 4. The relevant work is accepted for publication in the *International Journal of Production Research* (Eruguz et al., 2014).

Our final contribution concerns the mitigation of the CSL deviation in the GSM setting. Analysis provided for the bounded demand assumption in Chapter 3 shows that the effectively observed CSL at a final customer stage may be less than the target one. In the first part of this study, we show how to mitigate the CSL deviation by adjusting the safety factors applied at different stages of the supply chain. We propose two mitigation approaches and compare their performances in terms of total cost and computational time. The numerical study conducted on a real-world system shows that the first approach outperforms the second one in terms of computational time (30 seconds vs. 25 minutes) while the second approach provides better solutions in terms of cost (8.9% vs. 37.3% cost increase). In the second part of this study, we focus on decentralised systems where different parts of the supply chain are controlled by different actors. In such systems, even the CSL deviation is mitigated for each actor, the most downstream actors may still face a CSL deviation due to the demand bounds applied at the upstream actor. We show that the CSL deviation may be significant when the

real-world assembly system presented by Graves and Willems (2003) is controlled by two actors (8.1% on average). This contribution is presented in Chapter 5. A preliminary version of this study is accepted for publication in the proceedings of the *5th International Conference on Industrial Engineering and Systems Management*, *IESM'13* (Eruguz et al., 2013a) and is presented at the 2013 *INFORMS Manufacturing and Service Operations Management (MSOM) Conference* (Eruguz et al., 2013b).

Structure

This thesis is divided into 5 chapters. After giving a general introduction in this chapter, Chapter 1 outlines fundamentals that form the basis of the upcoming chapters. It includes the basic terminology and detailed presentation of the two main models, the GSM and the SSM that deal with the multi-echelon safety stock optimisation problem. Chapter 2 provides a comprehensive literature review of the multi-echelon safety stock optimisation approach considering both approaches. Chapter 3 presents the analysis conducted to quantify the impact of the bounded demand, guaranteed service times and common review periods assumptions of the GSM. Chapter 4 provides an extension of the GSM that enables to optimise the reorder intervals and order-up-to levels of stages simultaneously. It also presents detailed insights on an optimisation procedure proposed to find a reasonable solution the relevant optimisation problem. Chapter 5 deals with the mitigation of the CSL deviation under the GSM setting.

CHAPTER 1: PRELIMINARIES

The goal of this chapter is to provide the reader with the basic terminology that forms the basis of the upcoming chapters (Section 1.1) and to present the two main models, i.e. the Guaranteed-Service Model (GSM) and the Stochastic-Service Model (SSM) that deal with the safety stock optimisation problem in multi-echelon systems (Section 1.2). Although the main focus of this thesis is the GSM approach, both models are presented in order to be complete in terms of existing approaches dealing with the multi-echelon safety stock optimisation problem. The reader interested in the SSM approach can find further details in Axsäter (2006), Van Houtum (2006) and Simchi-Levi and Zhao (2012). The GSM presented in Section 1.2.2 corresponds to the original model on which this thesis' contributions are build.

1.1 Basic Terminology

This section introduces the basic terminology, definitions and notations relative to: 1) system structure, 2) demand, 3) lead time, 4) cost components, 5) service measures, 6) stock components and 7) inventory control policies used in multi-echelon inventory optimisation models.

1.1.1 System Structure

In a supply chain system, each stage is associated with certain processes such as the procurement of raw materials, the manufacturing or the transportation of items. Besides, each stage is considered as a potential location for holding the stock of the item processed at this stage.

A single-stage supply chain may typically be represented as a single company where input items are replenished from a supplier, processed within the company and then put into stock to satisfy a customer demand (Figure 1.1). Such systems have been the primary focus of inventory systems and have widely been studied to date (see, e.g., Silver et al., 1998; Zipkin, 2000).

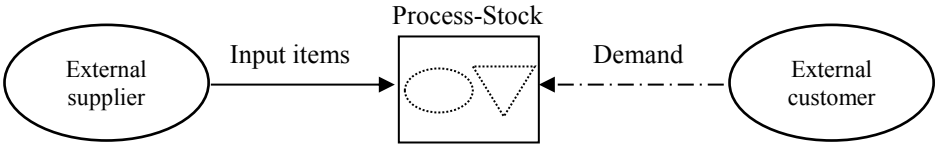

Figure 1.1: Single-stage system

In reality, supply chains usually consist of multiple stages and can be characterised as multi-echelon systems. A multi-echelon system can be modelled as a network where nodes represent stages and directed arcs denote the precedence relationships between stages. There exists a directed arc form upstream to downstream between two nodes in the network if an upstream stage (a predecessor, an internal supplier) directly supplies a downstream stage (a successor, an internal customer). The number of echelons in such systems is the highest number of nodes on a path between a most upstream node and a most downstream node.

Multi-echelon systems can be classified according to their network structures. In a serial system (Figure 1.2a), each stage has a single successor and a single predecessor, in an assembly system (Figure 1.2b), each stage has at most one successor and in a distribution system (Figure 1.2c), each stage has at most one predecessor. Real-world supply chains usually represent general multi-echelon systems, i.e. combinations of assembly and distribution systems. According to data provided by Willems (2008), most of real-world supply chains in industries such as computer hard-ware, semiconductor, industrial chemicals, consumer goods and aircraft engine represent such structures. We classify general multi-echelon systems into two categories, general acyclic systems (Figure 1.2d) and general cyclic systems (Figure 1.2e). In general cyclic systems, cycles may represent the returns of used and disassembled items to the system. Another example is in chemical or pharmaceutical industries where some products are generated together with their outcomes.

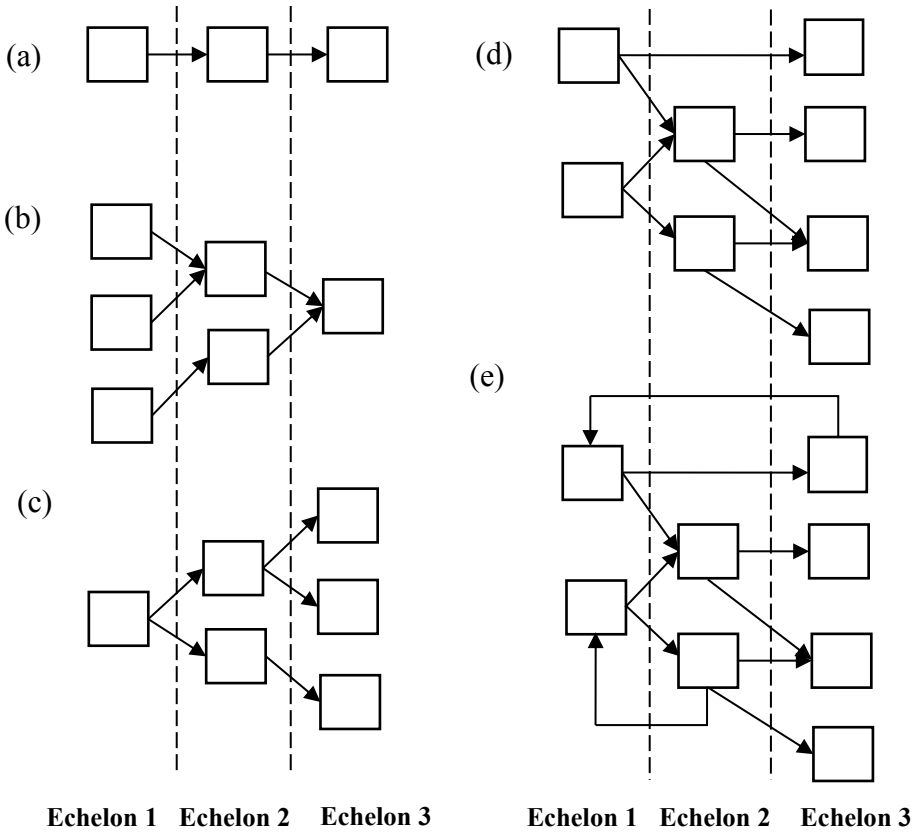

Figure 1.2: Network structures for multi-echelon systems

In a multi-echelon network, we denote the set of nodes by $\mathbf{N}$ and the set of arcs by $\mathbf{A}$. We define a scalar θ_{ij} associated with each couple of stages (i, j) that represents the number of input items required from upstream stage i to obtain one output item at downstream stage j if there exists a directed path or a directed arc between nodes i and j in the network. Besides, we partition the set of nodes (stages) into three disjoint sets: the set of supply nodes (supply stages) $\mathbf{N_S}$, the set of internal nodes (internal stages) $\mathbf{N_I}$ and the set of demand nodes (demand stages) $\mathbf{N_D}$. The set of supply nodes $\mathbf{N_S}$ is the set of nodes without predecessors, the set of demand nodes $\mathbf{N_D}$ is the set of nodes without successors and the set of internal nodes $\mathbf{N_I}$ is the set of nodes having at least one predecessor and one successor. For each node $j \notin \mathbf{N_D}$, we denote $\mathbf{N_D}(j)$ as the set of demand nodes which are connected to node j with a directed arc or path in the network. In other words, $\mathbf{N_D}(j)$ represents the set of demand stages that require the item processed at stage j.

1.1.2 Demand

In a single-stage system, demand comes from an external customer whereas in a multi-echelon system, processed items may be requested by external and/or internal customers. Internal demand at different stages of the supply chain can be derived based on external demand realisations and precedence relationship between various stages. In reality, such demand realisations might be uncertain. That is why for effective decision making in inventory management, one needs to forecast demands of future periods. For a single-stage system, several procedures are available for forecasting the short term future demand (see, e.g. Silver et al., 1998).

A common approach to model the uncertain demand is to assume a theoretical probability distribution function. Important parameters such as the mean and standard deviation of demand for a given period can be derived using the available data on previous demand realisations. Besides, one can use the information on forecasts and forecast errors while determining these parameters (see Babai, 2005). In the literature, items are classified as fast and slow moving items according to the size of their demand (Silver et al., 1998). For slow-moving items, demand process is often defined considering the inter-arrival process of customers and the distribution of their order size. For these items, demand is usually modelled as a (compound) Poisson process. For fast moving items, one can assume a theoretical probability distribution for the cumulative quantity of items requested within a single time period. Typically, a Normal distribution can be used to model the uncertain demand in this case. It is also possible to use the empirical demand distribution given by the available data. However, the use of the theoretical approach often enables the derivation of solution properties for the relevant optimisation problems.

In multi-echelon systems, we assume that the external demand is propagated to upstream stages, i.e. the internal demand at a stage can be derived based on the demands that occur at its downstream stages. For a given theoretical probability distribution for the external demand at stage $j \in \mathbf{N_D}$, we denote the mean period demand by μ_j, its standard deviation by σ_j and the correlation coefficient between the demands that occur at stages $j \in \mathbf{N_D}$ and $k \in \mathbf{N_D}$ by ρ_{jk}. If the correlation coefficient ρ_{jk} between the demands of two final items

associated with demand stages $j \in \mathbf{N_D}$ and $k \in \mathbf{N_D}$ is positive then, the demand of one item induces a demand for the other item. If ρ_{jk} is negative then, one item can be used as a substitute for the other one. The correlation coefficient equals zero for two items with independent demand. When external demand information is directly transmitted to all connected stages and no correlation exists between demands of different time periods, the mean and standard deviation of internal demands at non-demand stages may be expressed by the following equations:

$$\mu_i = \sum_{j:(i,j)\in\mathbf{A}} \theta_{ij}\mu_j \qquad \text{for } i \in \mathbf{N_I} \cup \mathbf{N_S} \tag{1.1}$$

$$\sigma_i = \sqrt{\sum_{j\in\mathbf{N_D}(i)} \sum_{k\in\mathbf{N_D}(i)} \theta_{ij}\theta_{ik}\sigma_j\sigma_k\rho_{jk}} \qquad \text{for } i \in \mathbf{N_I} \cup \mathbf{N_S} \tag{1.2}$$

One can assume that the internal demand at non-demand stage $i \in \mathbf{N_I} \cup \mathbf{N_S}$ follows the same type of theoretical probability distribution than the associated external demands and use (1.1) and (1.2), respectively, to estimate the mean and standard deviation of the internal demand at this stage (see, e.g., Minner, 1997; Inderfurth and Minner, 1998).

1.1.3 Lead Time

The lead time of a stage represents the duration of the process being realised at this stage, given that all necessary input items are available to start the process. The lead time may include several components regarding the stage process such as order processing, transportation, waiting, manufacturing, packing and storing times. Order processing time represents the time required for administrative processes at the relevant stage or at its external or internal suppliers. The transportation time is the time required to transport all the input items from the external/internal suppliers. The waiting time occurs when items use the same resources for the execution of the stage process and should wait until these resources become available. The manufacturing time is the duration of all manufacturing operations. The packing and the storing time is the time that elapses to put the processed item into inventory. The lead time ends when items become available for internal/external customer demand. As defined here, we do not include the transportation time to customers into the lead time since

this is considered as a part of the lead time of customers. The lead time L_j of stage $j \in \mathbf{N}$ may be characterised as stochastic due to several uncertainties regarding the nature of processes being realised (e.g., machine breakdowns, occupation of resources, congestion, order processing time variations etc.).

Under demand uncertainty, there is a chance of not being able to satisfy some of demand immediately. If demand is larger than the amount of stock held at a stage, a stock-out occurs. In case of backordering, the customers accept to wait until the out-of-stock items become available. In case of lost sales, the unfulfilled demand is lost. When demand is uncertain and backorders are allowed, an additional time component to consider is the waiting time of stages that stems from the delay in obtaining the backordered items. The replenishment time $\tilde{L}_j$ of stage $j \in \mathbf{N}$ is the sum of its lead time L_j and its waiting time due to this stock-out delay. We note that under demand uncertainty, due to occasional stock-outs at external/internal suppliers, the replenishment time of a stage becomes stochastic event if its lead time is deterministic.

1.1.4 Cost Components

There exist several cost components to consider when managing inventories in supply chains. These include: stage cost, processed item cost, holding cost, fixed ordering cost and shortage cost.

The stage cost c_j at stage $j \in \mathbf{N}$ is the per-unit cost that stems from the stage process. For instance, at stage $j \in \mathbf{N_S}$, the stage cost represents the unit purchasing price paid to the external supplier plus any cost incurred to make the item available to serve the demand of internal customers. For all other stages $j \in \mathbf{N_I} \cup \mathbf{N_D}$, the stage cost is the cost incurred to realise the relevant process at this stage. It can eventually include the mark-up applied by the upstream echelon if the considered supply chain involves different companies. The cost of a processed item p_j at stage $j \in \mathbf{N_I} \cup \mathbf{N_D}$ can then be defined as the total amount of money that has been spent to make the item available for usage of internal/external customers. Hence, the cost of a processed item p_j at stage $j \in \mathbf{N_I} \cup \mathbf{N_D}$ can be computed by adding the cost of all input items involved in the process to c_j:

$$p_j = c_j \qquad \forall j \in \mathbf{N_S}$$
$$p_j = c_j + \sum_{i:(i,j)\in\mathbf{A}} \theta_{ij} p_i \qquad \forall j \in \mathbf{N_I} \cup \mathbf{N_D}$$

The holding cost rate υ is the cost of holding one monetary unit (such as one euro, dollar, yen etc.) of inventory per period. This is usually made up of the opportunity cost of money invested, the costs of handling, special storage requirements, damage, insurance and possibly taxes. To make the inventory decision more manageable, usually, a single value of υ is assumed for all items (Silver et al., 1998) and the following rule is used to derive the per-unit holding cost per period h_j at stage *j*:

$$h_j = p_j \upsilon \qquad \forall j \in \mathbf{N}$$

Under this setting, the per-unit echelon holding cost per period h_j^e at stage *j* can be defined as:

$$h_j^e = c_j \upsilon \qquad \forall j \in \mathbf{N}$$

The fixed ordering cost A_j is the cost incurred each time input items are ordered from external/internal suppliers by stage $j \in \mathbf{N}$. It may consist of the setup cost incurred in manufacturing and/or other costs such as the cost of order forms, telephone calls, receiving and inspection of items etc. We note that the variable ordering cost is included into the stage cost.

There exist several ways of costing a stock-out at a stage. Among well-known shortage cost measures, the first one is to apply a fixed shortage cost for each stock-out occasion. Hence, the shortage cost at a stage becomes independent of the magnitude or the duration of stock-outs. The second measure corresponds to a shortage cost that is charged for each unit of out-of-stock item. In this case, the shortage cost is dependent of the magnitude of stock-outs and independent of their duration. The third measure is defined as a shortage cost that is incurred per each unit of out-of-stock item and per period. This considers both the magnitude and the duration of stock-outs. We note that the first and second measures can be applied in either backordering or lost sales settings whereas the third measure can only be applied in case of backordering.

In multi-echelon systems, shortage costs may be considered as given input parameters for stages facing external customer demand. However, in practice, choosing the appropriate shortage cost measure, estimating the unit-cost value and the expected magnitude of stock-outs are not easy tasks. If such information is available within an inventory optimisation context, one can aggregate all cost components and look for a solution that minimises the total cost. If this information is not available, an alternative formulation can be developed considering the customer service level constraints. The customer service level of a stage refers to the service provided to its customers in fulfilling their demand. The following section presents well-known service measures referred in the literature while defining the customer service level constraints.

1.1.5 Service Measures

In multi-echelon systems, we distinguish between two types of service performances: internal and external service performances. The latter are related to the service level provided to external customers whereas the former is related to internal customer service. Diks et al. (1996) emphasise that in multi-echelon systems, the internal service levels may be irrelevant as long as the external service levels reach the exogenously specified targets at the lowest cost. There exist two types of service measures frequently used in the literature, the Cycle-Service-Level (CSL) and the Fill-Rate (FR).

The CSL α_j at stage $j \in \mathbf{N}$ is defined as the non-stock-out probability during a cycle at this stage. A cycle may be defined as the time between two successive replenishments. In this case, the CSL is the fraction of replenishment cycles in which a stock-out does not occur.

The FR β_j is the fraction of customer demand that is met without backorders or lost sales:

$$\beta_j = 1 - \frac{\text{mean unsatisfied demand per unit of time}}{\text{mean demand per unit of time}}$$

The computation of the FR service level is not always straightforward (Minner, 2000). An exact method and two approximations are provided by Zhang and Zhang (2007) for a single-stage system facing normally distributed demand.

For ease of computation, several researchers consider a modified fill-rate service measure γ_j (see, e.g., Johnson et al., 1995; Silver et al., 1998; Silver and Bischak, 2011):

$$\gamma_j = 1 - \frac{\text{mean unsatisfied demand before arrival of a replenishment order}}{\text{mean demand during an arbitrary replenishment cycle}}$$

Schneider (1981) and Minner (2000) introduce another definition for the modified fill-rate service measure γ'_j:

$$\gamma'_j = 1 - \frac{\text{mean cumulative unsatisfied demand per unit of time}}{\text{mean demand per unit of time}}$$

The difference between these two definitions is that γ_j is based on the behaviour of the stock at the end of a replenishment cycle whereas γ'_j considers the behaviour per time unit. We note that for high service levels, i.e. as long as demand is very rarely backordered for more than one period the FR and modified fill-rate service measures are almost identical.

The choice of the service measure to use and the determination of the target service levels are decisions that are based on managerial experience and company strategy. Silver et al. (1998) summarise the factors that influence such decisions. Indeed, the decision can differ from item to item. Several factors such as market competition, customer preferences, their behaviour in stock-out situations and the availability of measures to resort in case of stock-outs influence the relevant decisions.

1.1.6 Stock Components

Stocks in supply chains have several components that can be classified regarding their motives or the constraints from which they arise.

If ordering decisions cannot be performed continuously and are placed at certain points of time, items should be ordered in batches. In this case, the stock level of a stage reaches an upper level just after the arrival of a batch and a lower level just before the arrival of the next batch and so forth. Cycle stocks stem from these cycles. The reasons that induce cycle stocks may be the existence of fixed ordering costs (i.e. the economies of scale) or the incapacity of an information system to continuously monitor the stock status and to place orders.

The existence of lead times causes another stock component that is called the pipeline stock. The pipeline stock at a stage includes all items that are in process at this stage as well as those are in transit to this stage. The level of pipeline stock at a stage depends on its lead time and mean demand, i.e. on parameters that are often considered as input parameters for the inventory decision problems. Therefore, pipeline stocks are usually neglected in the considered optimisation problems since they do not affect inventory decisions.

Safety stock is introduced as a lever against demand and replenishment time uncertainties in supply chains. The safety stock at a stage refers to the expected stock level at this stage just before an order arrives. Safety stocks are required to remedy stock-outs for situations where what is received deviates from what is delivered in quantity and time (Hax and Candea, 1984).

In practice, there exist other motives to hold stock in supply chains such as speculation and anticipation. The speculation stock may stem from an expected price increase in purchased items from external suppliers. The anticipation stock may be induced by a time varying demand pattern (seasonality), rather than expectations.

1.1.7 Inventory Control Policies

A number of possible inventory control policies are introduced for single-stage systems. These policies are classified into two major categories regarding how the inventory status is reviewed: continuous-review policies and periodic-review policies.

In continuous-review policies, the stock status is continuously monitored and an order to replenish items is placed immediately after the stock position of the stage (the sum of all its planned orders and its physical stock minus its backorders) drops below a reorder point r. If each order size is equal to a fixed quantity Q then the relevant policy is called the order-point, order-quantity (r, Q) policy. Another popular policy in this category is the order-point, order-up-to level policy (r, S) in which the order sizes are such that the stock position at the stage returns to a target order-up-to level S just after each order placement. An important special case is when $r = S - 1$ where the policy is called the continuous-review, base-stock $(S-1, S)$ policy.

In periodic-review policies, the stock status is inspected every R units of time and a replenishment order can be placed only at these review instants. A popular policy in this category is the periodic-review, order-up-to (R, S) policy for which the control procedure is to raise the inventory position to the order-up-to level S at each review instant R. When the review period $R = 1$, the relevant policy is also called as the periodic-review, base-stock policy. Furthermore, (R, r, S) and (R, r, Q) policies can be seen as periodic-review analogues of (r, S) and (r, Q) policies, respectively.

Inventory control policies presented above are also applicable for multi-echelon systems. However, in multi-echelon systems, the inventory control can be executed in two manners, using the installation stock or the echelon stock information. The use of installation stock information (installation stock policies) leads to a decentralised (local) control in the sense that ordering decision at a stage is only based on the inventory position of this stage. In this case, available information contains only the locally available inventory status. As a consequence, excessive demand may not be identified at upstream stages due to the delay in information through the considered ordering policy. The shortcoming of using such local information is avoided by echelon stock policies, i.e. by controlling the inventory based on the echelon inventory position of a stage. The echelon stock of a stage can be defined as the stock on hand at this stage plus all stock in the downstream part minus the backorders at the most downstream stages. The echelon inventory position of a stage is the sum of all stock in process at or in transit to this stage plus its echelon stock. The echelon stock concept is first introduced by Clark and Scarf (1960). Echelon stock policies require a complete knowledge of how much stock is at downstream stages and hence, necessitates an appropriate information technology.

We note that a stage that uses an installation stock policy can always raise its stock position to a desired level. Hence, under the backordering assumption, one can model each stage of the supply chain as a single stage system with random replenishment time. In echelon stock policies, the echelon stock position of a stage includes the stock in process at or in transit to this stage but does not include the upstream backorders. For these policies, the replenishment time of a stage equals to its lead time. However, it is more difficult to

determine the echelon stock position, since different stages cannot be regarded as single stage systems (Chen and Zheng, 1994a). Axsäter and Rosling (1993) prove that when each stage in a supply chain is controlled by a base-stock policy, an installation stock policy can always be replaced by an echelon stock policy and vice versa. When each stage is controlled by an order-point, order-quantity policy an installation stock can always be replaced by an echelon stock policy, but not vice versa. Indeed, echelon stock policies are superior to installation stock policies in terms of cost for the latter case (Axsater and Juntti, 1996; Axsäter and Juntti, 1997). For divergent systems the examples for the superiority of both policies can be found (Axsäter, 1997; Axsäter and Juntti, 1997).

1.2 Safety Stock Optimisation in Multi-Echelon Systems

The objective of safety stock optimisation problem analysed in either single or multi-echelon systems is to balance two types of risks, the risk of significant stock-outs for external customers and the risk of holding unnecessarily large inventory in the supply chain. The safety stock optimisation problem can be formulated either by introducing shortage costs or service level constraints. In the first case, one should specify a way of costing stock-outs and then search for a solution that minimises the total cost of shortage and holding inventory in the supply chain. In the second case, the objective is to minimise the total cost of holding inventory subject to target customer service levels (see Section 1.1.5). Indeed, these two approaches may lead to equivalent formulations for the considered single (see Silver et al., 1998) or multi-echelon (see Minner, 2000; Van Houtum and Zijm, 2000) systems. In practice, the service level approach may be preferred to the shortage cost approach since shortage costs are often hard to estimate. Besides, even in situations where shortage costs are explicitly defined (e.g. when there exist contractual specifications regarding the stock-out penalties), service measures are still needed to track the service level performance of the supply chain (Diks et al., 1996).

The use of a multi-echelon approach for safety stock optimisation provides better results in terms of cost and customer service level performance in comparison to the single-echelon approach that consists of determining independently the local safety stock associated

with each stage in the supply chain. However, the use of a multi-echelon approach is a difficult task due to the number of interdependent decision variables and non-linear functions. The complexity of this approach is directly related to the considered network structure since it depends on the number of stages and the topology of stage connections in the network. Indeed, multi-echelon safety stock optimisation represents a computational challenge especially for general networks, i.e. for most of real-world supply chains.

The multi-echelon safety stock optimisation problem is widely studied in the literature. In order to address this problem, researchers proposed the Stochastic-Service Model (SSM) and the Guaranteed-Service Model (GSM) approaches that are introduced by Clark and Scarf (1960) and Simpson (1958), respectively. These two approaches differ in terms of assumptions made with regard to how to address demand variations and service times. Briefly, in SSM approach, each stage in the supply chain provides an immediate service when stock is on hand but entails a stochastic delay in case of stock-out. The GSM assumes that after a certain service time which is quoted to the downstream stages, items are always available (guaranteed-service time assumption). This is achieved by establishing upper bounds for demand at each stage of the supply chain (bounded demand assumption). The GSM setting enables to consider real-world supply chains that are usually characterised as large and complex multi-echelon systems whereas the SSM approach mostly focuses on more simple and smaller supply chain structures such as serial, assembly or two-echelon distribution systems (Hwarng et al., 2005).

In what follows, we first present the most commonly known model pertaining to the SSM approach, i.e. the Clark and Scarf (1960) model (Section 1.2.1). Then, we provide a detailed presentation of the original GSM (Section 1.2.2).

1.2.1 Stochastic-Service Model

The origin of the SSM approach is the seminal work of Clark and Scarf (1960) who prove that the cost optimal inventory control policy for all stages of a serial system is an echelon order-up-to (base-stock) policy. Since this work, a lot of research has been done to consider different network structures and different assumptions concerning external demand

process, lead times, ordering policies etc. (Simchi-Levi and Zhao, 2012). However, within the SSM approach, a generic model that fits well with any type of supply chain network structure does not exist. The models and solution procedures proposed are much different according to the network structures and assumptions considered. We provide a literature review of these models in Chapter 2. For ease of presentation, we present in this section the model of Clark and Scarf (1960) for a two-stage serial system operating on an infinite time horizon. In the considered system, Stage 2 replenishes from Stage 1 and Stage 1 replenishes from an external supplier (see Figure 1.3). In what follows, we first summarise the assumptions of this model. Second, we present the inventory dynamics under the considered assumptions. Third, we give the solution method, i.e. the decomposition technique introduced by Clark and Scarf (1960). Finally, we provide a numerical example. For more details, we refer to Axsäter (2006) and Van Houtum (2006).

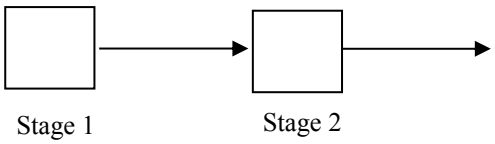

Figure 1.3: Two-stage serial system

Modelling Assumptions

The model considers the following assumptions:

(i) Time is divided into base planning periods, i.e. periods of equal length which can be days, weeks, months etc. An infinite time horizon is considered.

(ii) All stages are allowed to place orders at the beginning of each period (the review period equals one period).

(iii) The lead time of each stage is assumed to be deterministic, constant and an integer multiple of the base planning period. Lead time at the most downstream stage includes the review period of one period length.

(iv) There are no capacity constraints in the system regarding physical space or volume of work.

(v) The external supplier has infinite capacity.

(vi) A linear inventory holding cost structure is considered.

(vii) External demand occurs at the most downstream stage (demand stage) and demands in different periods are identically and independently distributed (*i.i.d.*) on $[0,\infty)$.

We note that the assumptions considered so far are also valid for the GSM. The main difference is due to assumptions concerning the external demand modelling. In the SSM that we present in this section, external demand can be modelled by a continuous demand distribution with an average μ and a standard deviation σ per period. Demand that cannot be met directly from stock is backordered. A per-unit shortage cost b_2 incurs at Stage 2 per backordered item and period. Besides, without loss of generality, the number of input items required from Stage 1 to obtain one output item at Stage 2 equals one.

The model presented here is expressed in terms of cost minimisation, i.e. the aim is to minimise the total expected holding and backorder costs per period. Indeed, an equivalent formulation under service level constraints can be obtained as will be presented in this section (page 25). We note that at this stage we do not assume anything about the inventory control policy to be applied for the considered system.

Inventory Dynamics

All events take place in each period in the following order: (1) an order is placed at each stage, (2) orders arrive, (3) demand occurs, (4) costs are evaluated. The first two events take place at the beginning of each period. The last event occurs at the end of each period. The third event, the demand, may occur anywhere in between for the demand stage. Since we assume a periodic-review, for non-demand stages, demand in each period occurs once at the beginning of the period. In any case, we assume that demand occurs after the second event, i.e. after the arrival of orders.

We consider that after ordering in period *t*, Stage 1 has a certain echelon inventory position y_1. Because the outside supplier has infinite supply, orders of Stage 1 are always satisfied without backorders. Consider then the echelon inventory level $I_1^e(t+L_1)$ at Stage 1 in period $t+L_1$ just before the period demand. We can express $I_1^e(t+L_1)$ as y_1 minus the stochastic demand during L_1 periods, $t,t+1,..,t+L_1-1$ denoted by $d[t,t+L_1-1]$:

$$I_1^e(t+L_1) = y_1 - d[t, t+L_1-1] \tag{1.3}$$

We note that $d[t, t+L_1-1]$ has mean μL_1 and standard deviation $\mu\sqrt{L_1}$.

Next, we consider that the echelon inventory position at Stage 2 after ordering in period $t+L_1$ is equal to some level y_2. By definition, the echelon inventory position at Stage 2 does not include the backorders at Stage 2 (i.e. the backorders at Stage 2 are deducted from its echelon inventory position). Here, whatever policy is followed, the following must be held:

$$y_2 \le I_1^e(t+L_1) = y_1 - d[t, t+L_1-1] \tag{1.4}$$

We use the notation $(x)^+ = \max\{x,0\}$ and $(x)^- = \max\{-x,0\}$. In case of strict inequality in (1.4), the difference represents the positive value of installation stock $I_1(t+L_1)$ at Stage 1 just after the order from Stage 2:

$$(I_1(t+L_1))^+ = I_1^e(t+L_1) - y_2 \tag{1.5}$$

Equation (1.5) gives the installation inventory level at Stage 1 at the end of period $t+L_1$.

We note that the lead time L_2 of Stage 2 includes the review period. The installation inventory level $I_2(t+L_1+L_2-1)$ at Stage 2 after the demand in period $t+L_1+L_2-1$ is obtained as y_2 minus the demand in periods $t+L_1, t+L_1+1, ..., t+L_1+L_2-1$:

$$I_2(t+L_1+L_2-1) = y_2 - d[t+L_1, t+L_1+L_2-1] \tag{1.6}$$

The installation inventory level at Stage 2 at the end of period $t+L_1+L_2-1$ becomes (1.6) where $d[t+L_1, t+L_1+L_2-1]$ has mean μL_2 and standard deviation $\mu\sqrt{L_2}$.

Solution Method

Here, we present the optimality of echelon order-up-to policies and the solution method called the decomposition technique to show how to determine the optimal echelon order-up-to levels $\hat{S}_1^*$ and $\hat{S}_2^*$ that minimise the total cost.

Using the results presented in the previous subsection, we can present the expected cost at Stage 1 at the end of period $t+L_1$ and the expected cost at Stage 2 at the end of period $t+L_1+L_2-1$. First, using (1.3) and (1.5), we can express the former as:

$$C_1 = h_1 E[I_1^e(t+L_1) - y_2] = h_1 E[y_1 - d[t, t+L_1-1] - y_2] \\ = h_1(y_1 - \mu L_1) - h_1 y_2 \tag{1.7}$$

Second, due to (1.6) the expected cost at Stage 2 at the end of period $t+L_1+L_2-1$ becomes:

$$C_2 = h_2 E[(y_2 - d[t+L_1, t+L_1+L_2-1])^+] + b_2 E[(y_2 - d[t+L_1, t+L_1+L_2-1])^-] \\ = h_2(y_2 - \mu L_2) + (h_2+b_2) E[(y_2 - d[t+L_1, t+L_1+L_2-1])^-]$$

If we transfer the last term in (1.7) to C_2 we obtain the following costs:

$$\tilde{C}_1 = h_1(y_1 - \mu L_1) \tag{1.8}$$

$$\tilde{C}_2 = h_2^e y_2 - h_2 \mu L_2 + (h_2+b_2) E[(y_2 - d[t+L_1, t+L_1+L_2-1])^-] \tag{1.9}$$

This reallocation does not affect the total cost. We observe (1.8) is independent of y_2. The cost expressed in (1.9) may depend on y_1 due to the relation in (1.4).

We note that periods $t+L_1$ and $t+L_1+L_2-1$ can be seen as arbitrary periods (Axsäter, 2006). Let ignore the dependency between $\tilde{C}_2$ and y_1 and assume that we can choose any value $\hat{S}_2$ for y_2. Hence, we can replace (1.9) by:

$$\hat{C}_2(\hat{S}_2) = h_2^e \hat{S}_2 - h_2 \mu L_2 + (h_2+b_2) \int_{\hat{S}_2}^{\infty} (u - \hat{S}_2) f_{L_2}(u) du \tag{1.10}$$

where $f_L(u)$ is the L-period demand probability density function. It is easy to obtain the optimal value $\hat{S}_2^*$ from the first order condition:

$$\int_0^{\hat{S}_2^*} f_{L_2}(u) du = \frac{b_2 + h_1}{b_2 + h_2} \tag{1.11}$$

If $\hat{S}_2^*$ obtained by solving (1.11) is such that $\hat{S}_2^* \le y_1 - d[t, t+L_1-1]$, then $\hat{S}_2^*$ is the optimal solution. But if $\hat{S}_2^* > y_1 - d[t, t+L_1-1]$, then the best possible value of y_2 is $y_2 = y_1 - d[t, t+L_1-1]$ due to the convexity of (1.10). The optimal policy can be realised if

Stage 2 operates with an echelon order-up-to policy where its echelon order-up-to level is equal to $\hat{S}_2^*$. Indeed, there is no difference between echelon stock and installation stock at Stage 2.

If we assume that Stage 2 uses the optimal echelon order-up-to level $\hat{S}_2^*$, for a given level $\hat{S}_1$ of y_1, we can rewrite the total cost of the two-stage serial system using (1.8) and (1.10):

$$\hat{C}_1(\hat{S}_1) = h_1(\hat{S}_1 - \mu L_1) + \hat{C}_2(\hat{S}_2^*) + \int_{\hat{S}_1 - \hat{S}_2^*}^{\infty} \left[\hat{C}_2(\hat{S}_1 - u) - \hat{C}_2(\hat{S}_2^*)\right] f_{L_1}(u)du \qquad (1.12)$$

The last term in (1.12) can be seen as the shortage cost at Stage 1 due to its inability to satisfy the demand of Stage 2 on time. It is easy to verify that (1.12) is convex. Hence, from the first order condition we can obtain the following:

$$\int_0^{\hat{S}_2^*} \int_0^{\hat{S}_1^* - u} f_{L_1}(v) f_{L_2}(u)dvdu = \frac{b_2}{b_2 + h_2} \qquad (1.13)$$

Since the external supplier has infinite supply it is optimal to apply an echelon order-up-to policy with echelon order-up-to level equal to $\hat{S}_1^*$ at Stage 1. Therefore, it is optimal to apply an echelon order-up-to policy for both stages of the considered system.

We note that the same problem can be modelled under a service level constraint. In particular, if there exist a predetermined CSL level α to be ensured towards the external customer, the cost parameter b_2 can be obtained from the following equivalence relation:

$$\alpha = \frac{b_2}{b_2 + h_2} \qquad (1.14)$$

As shown in this subsection, the optimal echelon order-op-to levels can be obtained by sequentially minimizing one-dimensional convex functions. The relevant procedure is called the decomposition technique. In the first step of this technique, the optimal order-up-to level of the demand stage is obtained. In the next step, the additional cost due to the stock-outs at the upstream stage is evaluated and added to the total cost when determining an optimal policy for the upstream stage. Under the assumptions considered in this section, the decomposition technique leads to the derivation of (1.11) and (1.13). These equations are

called the Newsboy equations as they constitute a generalisation of the well-known Newsboy equation developed for single-stage systems (see Hadley and Whitin, 1963). The decomposition technique is easy to generalise to serial systems with more than two stages (see Chapter 2, Section 2.1).

Numerical Example

The numerical example developed here is inspired from Axsäter (2006). We assume that the external demand at Stage 2 is continuous, normally distributed and independent across periods with parameters $\mu = 10$ and $\sigma = 5$ per period. Holding cost and lead time data of stages are as follows: $h_1 = 1$, $h_2 = 1.5$, $L_1 = 5$ and $L_2 = 6$ (including the review period at Stage 2). The per-unit backordering cost b_2 at Stage 2 is equal to 20.

Using the decomposition technique, optimal echelon order-up-to levels $\hat{S}_1^* = 135.7$ and $\hat{S}_2^* = 84.4$ can be obtained from Newsboy equations (1.11) and (1.13), respectively (we use numerical integration). This gives the total cost $C_{SSM} = 46.0$. We note that the installation order-up-to level of stages are $S_1^* = \hat{S}_1^* - \hat{S}_2^*$ and $S_2^* = \hat{S}_2^*$. The optimal solution obtained is illustrated in Figure 1.4.

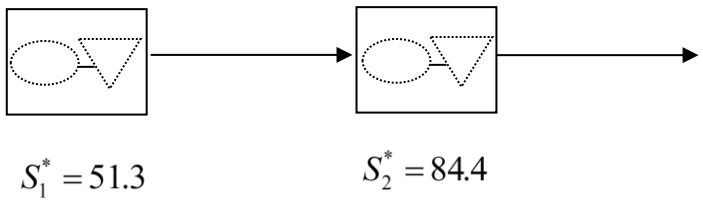

Figure 1.4: The solution obtained by the SSM approach

1.2.2 Guaranteed-Service Model

The GSM has been developed first for a single-stage inventory system by Kimball (1988), whose paper was originally written in 1955. Simpson (1958) extends this framework to a serial system and initiates the research on the GSM for multi-echelon systems. Graves and Willems (2000) generalise the GSM formulation to general multi-echelon systems. This section presents the modelling assumptions, inventory dynamics and mathematical programming formulation of the original GSM in order to introduce the model as it is usually

referred in the literature. We also provide a numerical example at the end of this section for illustration.

Modelling Assumptions

In the GSM, the multi-echelon system can be modelled as a network of any structure. The original GSM holds assumptions (i)-(vii) presented for the SSM in Section 1.2.1 (bearing in mind that in a general network there may exist several demand stages). The main difference of the GSM stems from the inventory policy, the external demand modelling and the service time notion.

In the GSM, it is assumed that each stage $j \in \mathbf{N}$ follows a periodic-review (R_j), installation order-up-to (S_j) policy. As a result of assumption (ii), all stages have the same review period. The length of the review period is equal to the length of the base planning period.

External demand occurs only at nodes without successors, i.e. at demand stages. For each demand stage $j \in \mathbf{N_D}$, the external demand comes from a stationary *i.i.d.* process with an average μ_j and a standard deviation σ_j per period. For demand stage $j \in \mathbf{N_D}$, we denote the demand during $(t-1,t]$ by $d_j(t)$.

Non-demand stages (internal or supply stages) have only internal customers that are their immediate downstream stages. Hence, the realised demand $d_i(t)$ at a non-demand stage $i \in \mathbf{N_S} \cup \mathbf{N_I}$ in period *t* is the sum of the orders placed by its immediate downstream stages. Under the inventory control policy considered, the realised demand $d_i(t)$ becomes:

$$d_i(t) = \sum_{j:(i,j)\in\mathbf{A}} \theta_{ij} d_j(t)$$

Therefore, internal demand parameters can be derived using (1.1) and (1.2).

Demand is assumed to be bounded with an increasing and concave function $D_j(\tau_j)$ for every stage $j \in \mathbf{N}$ and for any long period $\tau_j = 1,2,...,M_j$ where M_j is the maximum replenishment time, i.e. the maximum time that could elapse between the ordering and the reception of items at stage *j*. The maximum replenishment time at stage *j* can be calculated by:

$$M_j = L_j + \max\{M_i | i : (i, j) \in \mathbf{A}\}$$

Under this setting, for any period $t \geq \tau_j$ and for any long period $\tau_j = 1,2,...,M_j$ we assume the following for every stage $j \in \mathbf{N}$:

$$D_j(\tau_j) \geq d_j(t - \tau_j, t)$$

where $d_j(a,b)$ is the realised demand at stage *j* over the time interval $(a,b]$:

$$d_j(a,b) = 0 \qquad \text{for } a \geq b$$

$$d_j(a,b) = \sum_{t=a+1}^{b} d_j(t) \qquad \text{for } a < b$$

In practice, the existence of demand bounds does not imply that the arrival demand can never exceed the demand bounds. Instead, it reflects the maximum amount of demand that is satisfied from safety stock. When demand exceeds $D_j(\tau_j)$ over τ_j, the safety stock is regarded as not being adequate to satisfy the excess demand. In this case, the company might handle the excess demand by other countermeasures such as subcontracting, overtime production and/or express expediting. However, the original GSM does not explicitly model what happens in case of excessive demand and does not quantify the cost of such extraordinary measures. We note that in most of existing models, demand bounds are determined based on a target CSL which is specified according to the company's policy.

Furthermore, it is assumed that each stage $j \in \mathbf{N}$ promises an outbound guaranteed service time s_j^{out} to its downstream stages such that the realised demand $d_j(t)$ at stage *j* in period *t*, is totally satisfied (with 100% service) at period $t + s_j^{out}$. For all inbound arcs of stage *j*, the inbound service time s_j^{in} defines the time necessary for stage *j* to get all of its inputs from stages $i : (i, j) \in \mathbf{A}$ and to start its process. Since stage *j* cannot start the process without receiving all inputs, $s_j^{in} \geq s_i^{out}$ should be ensured for all arcs $(i, j) \in \mathbf{A}$. We note that outbound and inbound service times are decision variables of the GSM. Besides, the service times are considered as integer multiples of a base planning period. Without loss of generalisation, we consider the base planning period as one unit of time which makes $R_j = 1, \forall j \in \mathbf{N}$. Hence, the decision variables can be considered as positive integers.

The assumptions regarding demand bounds and services times differentiate the GSM from the SSM literature. These assumptions are controversial since they can be judged as restrictive and oversimplified. Nevertheless, they are crucial to guarantee a tractable model. Indeed, the bounded demand assumption becomes realistic when a manager is able to explicitly indicate a preference for the demand range to be covered by safety stocks and when the supply chain has sufficient operational flexibility to apply extraordinary measures in case of excessive demand. Usually, specifying an appropriate level for the demand range that would be covered by safety stocks is not straightforward since it requires a good managerial experience. In some contexts external demand can be bounded in the real sense due to the capacity constraints of external customers.

Inventory Dynamics

Under the assumptions presented in the previous subsection, at the beginning of each period *t*, the realised demand $d_j(t)$ at stage *j* is observed and an order corresponding to this demand is placed. The corresponding order is replenished (and available to serve demand) at the beginning of period $t + s_j^{in} + L_j$. On the other hand, Stage *j* is subject to satisfy this demand at period $t + s_j^{out}$. If demand $d_j(t)$ is served first and the replenishment corresponding to this demand occurs at a subsequent period, node *j* has to store the inventory that would satisfy this demand. That is, if $s_j^{in} + L_j > s_j^{out}$, stage *j* should have the amount of inventory to cover the demand over an interval of length $\tau_j = s_j^{in} + L_j - s_j^{out}$, that is called the "net replenishment time" of stage *j*.

We assume that the realised demand $d_j(t)$ in period *t* at stage *j* is equal to 0 for $t \le 0$ and the inventory system starts at time 0 with initial inventory level $I_j(0) = S_j \ge 0$. Graves and Willems (2000) provide the following balance equation for the net (installation) inventory level $I_j(t)$ at stage *j* at the end of period *t*:

$$I_j(t) = S_j - d_j(t - s_j^{in} - L_j, t - s_j^{out}) \tag{1.15}$$

The first argument of $d_j(.,.)$, $t - s_j^{in} - L_j$ corresponds to the last replenishment received by stage *j* by time *t*. The second argument of $d_j(.,.)$, $t - s_j^{out}$ corresponds to the last demand served by stage *j* by time *t*. At each time, the inventory is exposed to $s_j^{in} + L_j - s_j^{out}$

periods of demand which is its net replenishment time. In order to achieve guaranteed-service with minimum inventory at stage *j*, the order-up-to level S_j should be set equal to the demand bound $D_j(\tau_j)$ where $\tau_j = s_j^{in} + L_j - s_j^{out}$. In this way, the net inventory level $I_j(t)$ at stage *j* will always be positive and 100% service will be guaranteed to internal/external customers within the specified service times. Under this setting, the expected inventory level $E[I_j(t)]$ can be expressed as:

$$E[I_j(t)] = D_j(\tau_j) - \tau_j \mu_j \tag{1.16}$$

The expected inventory level expressed in (1.16) represents the safety stock held at stage *j*. As given by assumption (vi), a linear inventory cost structure is considered.

Mathematical Programming Formulation

The GSM optimisation problem is the problem of finding optimal outbound and inbound service times that minimise the total safety stock cost of the system:

$$P0: \min \sum_{j \in \mathbf{N}} h_j \left\{ D_j(\tau_j) - \tau_j \mu_j \right\} \tag{1.17}$$

$$\text{s.t.} \quad \tau_j = s_j^{in} + L_j - s_j^{out} \qquad \forall j \in \mathbf{N} \tag{1.18}$$

$$s_i^{out} \le s_j^{in}, \qquad \forall (i, j) \in \mathbf{A} \tag{1.19}$$

$$s_j^{out} \le s_j^{client}, \qquad \forall j \in \mathbf{N_D} \tag{1.20}$$

$$\tau_j, s_j^{out}, s_j^{in} \ge 0 \text{ and integer} \quad \forall j \in \mathbf{N} \tag{1.21}$$

The objective function (1.17) represents the minimisation of the total safety stock cost. Constraint (1.18) defines the net replenishment time of each stage. Constraint (1.19) ensures that the inbound service time of a stage is no smaller than the outbound service times of its upstream stages. Constraint (1.20) ensures that demand stages satisfy their service guarantee where s_j^{client} corresponds to the maximum service time that stage $j \in \mathbf{N_D}$ can quote for the external customer. Finally, with (1.21) all decision variables are restricted to be positive integers. The problem *P*0 is actually the minimisation of a concave function over a

closed, bounded convex set. After having solved the problem *P*0, the optimal order-up-to level S_j^* of stage j can be obtained by $S_j^* = D_j(\tau_j^*)$ where τ_j^* is the optimal net replenishment time of stage j.

Numerical Example

Let consider the example presented in Section 1.2.1 and use the GSM approach to determine the optimal service times and the associated installation order-up-to levels. To do this, first, we should specify the demand bound function. The setting proposed by Simpson (1958) and Graves and Willems (2000) lead to the following expression for the considered example:

$$D(\tau_j) = \mu\tau_j + z\sigma\sqrt{\tau_j} \text{ for } j \in \{1,2\} \tag{1.22}$$

where z is the safety factor that relates to a CSL α at Stage 2. For the considered example, we have $z = \Phi^{-1}(\alpha)$ where $\Phi(.)$ denotes the standard normal cumulative distribution function. We can obtain the CSL level $\alpha = 93.0\%$ using the equivalence relation given in (1.14). We assume that $s_j^{client} = 0$, i.e. external demand at Stage 2 is immediately satisfied when it occurs. The optimal service times and net replenishment times are obtained solving the problem *P*0: $s_1^{out} = s_2^{in} = 5,\ s_2^{out} = s_2^{client} = 0,\ \tau_1^* = 0$ and $\tau_2^* = 11.$ Hence, the optimal installation order-up-to levels are: $S_1^* = 0$ and $S_2^* = 134.5$. This solution gives the total cost $C_{GSM} = 36.8$. Figure 1.5 illustrates the solution obtained by the GSM.

The first difference between the solutions obtained using the SSM and GSM approaches is that in the SSM solution, both stages hold safety stock whereas in the GSM solution only Stage 2 holds safety stock. That is, Stage 1 delays the fulfilment of Stage 2's orders so that each order is met when the replenishment associated with this order is received. Indeed, in GSM solutions, only stages with strictly positive net replenishment times need to hold safety stock. Due to this feature, the GSM optimisation problem is also called the safety stock placement problem. The second difference between the two solutions stems from the bounded demand assumption. In the GSM, safety stocks are dimensioned to be protected against demand variability up to the specified demand bounds. The model does not quantify the cost of other countermeasures such as subcontracting, overtime production and/or express

expediting to handle the excess demand. In contrast, in the SSM, safety stock is the only countermeasure to use for all demand realisations. Indeed, the GSM underestimates the total cost of the supply chain since it does not model what happens in case of excessive demand. This explains the cost gap between C_{GSM} and C_{SSM}. In order to choose the best approach for the considered system, C_{GSM} should be considered together with the cost of extraordinary measures. It makes financial sense to adopt the GSM approach if the system has sufficient flexibility to address extraordinary measures in case of excessive demand and if the cost of these measures would be less than the cost gap $C_{SSM} - C_{GSM}$.

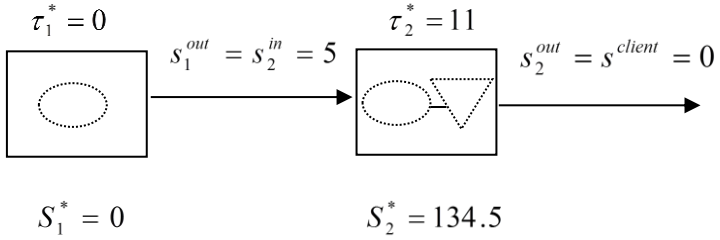

Figure 1.5: The solution obtained by the GSM approach

1.3 Conclusion

In this chapter, we have presented the basic terminology and the main models proposed for multi-echelon safety stock optimisation problem, the Clark and Scarf (1960) model which has initiated the research on SSM approach and the original GSM which has recently triggered several extensions. In the next chapter, we provide a literature review of the SSM and GSM approaches. We remind the reader that in this thesis our main contributions will concern the GSM approach.

CHAPTER 2: LITERATURE REVIEW

As said earlier, this thesis develops contributions to the multi-echelon inventory optimisation problem with a special focus on the GSM approach. In the literature, there exist two main approaches that enable to deal with the multi-echelon safety stock optimisation problem, the SSM approach and the GSM approach. This chapter presents a complete literature review considering both approaches.

The literatures of the SSM and GSM approaches are not at the same degree of maturity. The literature pertaining to the SSM approach is extensive. A lot of research has been done since the work of Clark and Scarf (1960). Excellent surveys are also a part of this vast literature (see, e.g., Diks et al., 1996; Axsäter, 2003; Simchi-Levi and Zhao, 2012). On the other hand, the research on GSM has gained interest only in the last decade. To the best of our knowledge, there is no review which gives a synthesis of the various works developed so far. The review conducted in this section enables us to reveal the gaps in the GSM literature that lead to the research questions considered in this thesis.

In Section 3.1, we aim at providing a review of the main models pertaining to the SSM approach. The purpose of Section 3.2 is to conduct a comprehensive review of the GSM approach. Section 2.3 then summarises the contributions that compare, contrast or combine the two approaches. Finally, Section 2.4 draws some conclusions by presenting the limitations of both approaches and presents the gaps in the GSM literature that are being addressed in the upcoming chapters of this thesis.

2.1 Stochastic-Service Model Approach

We observe that the models pertaining to the SSM approach mainly differ in terms of the multi-echelon system structure considered. Each model focuses on a specific system structure and aims at providing appropriate solutions under different assumptions. In this literature review, we classify the models pertaining to the SSM approach according to the system structure studied. The literature review is conducted over three axes: the models that study serial and assembly systems (Section 2.1.1), distribution systems (Section 2.1.2) and general systems (Section 2.1.3). We provide more details on the models that are developed

for general systems which are not comprehensively presented in the existing surveys of the SSM approach.

2.1.1 Serial and Assembly Systems

We remind that the seminal work of Clark and Scarf (1960) proves that the cost optimal inventory control policy for all stages of a serial system is an echelon order-up-to policy. Federgruen and Zipkin (1984b) extend this work to the infinite time horizon case. For this case, easier proofs for the optimality of echelon stock order-up-to policies are provided by Langenhoff and Zijm (1990), Chen and Zheng (1994b) and Van Houtum (2006). As presented in Section 1.2.1 (Chapter 1), the decomposition technique developed by Clark and Scarf (1960) leads to the derivation of Newsboy equations to determine the optimal echelon order-up-to levels under some assumptions. The Newsboy equations can generally not be solved analytically. However, computational procedures can be developed to determine the optimal echelon order-up-to levels. For a two-stage system and normally distributed demand, the optimal solution is relatively easy to obtain (see Federgruen and Zipkin, 1984a; 1984b). For the general serial case with number of stages greater than 2, the computations may be very time consuming. Van Houtum and Zijm (1991) provide approximate procedures among which one is exact for mixed Erlang demand distributions. Though these procedures employs the shortage cost approach it can alternatively be applied under the service level approach (Minner, 2000; Van Houtum and Zijm, 2000). We refer the reader to Van Houtum (2006) for a comprehensive review on existing exact and approximate procedures that enable to obtain echelon order-up-to levels and the relevant costs.

There exist several extensions/generalisations of the Clark and Scarf (1960) model. This includes the consideration of continuous-review, fixed order quantities (Bodt and Graves, 1985; Chen, 2000), capacity constraints (Parker and Kapuscinski, 2004), fixed reorder intervals (Van Houtum et al., 2007) and lost sales case (Huh and Janakiraman, 2010). Furthermore, Rosling (1989) and Langenhoff and Zijm (1990) demonstrate that an assembly system can be replaced by an equivalent serial system. Therefore, results presented for the

serial system can also be applied to assembly systems. We refer to the survey paper of Van Houtum et al. (1996) for an analysis of assembly systems.

2.1.2 Distribution Systems

It is common to use the Clark and Scarf (1960) approach also for distribution systems. However, in contrast to serial and assembly systems, additional problems stem from the requirement to allocate the available amount of stock among downstream stages in case of stock insufficiency at upstream stages. For distribution systems, the optimal echelon order-up-to policies can be derived under the so-called balance assumption. The balance assumption means that an upstream stage is allowed to make negative allocations to its downstream stages. In other words, the total stock at the downstream echelon can be optimally distributed between the associated stages in any period. Under this assumption, the optimal echelon order-up-to levels satisfy the Newsboy equations provided by Diks and De Kok (1998). Indeed, the determination of optimal allocation functions would be computationally infeasible for realistic problem instances due to its non-linearity. Researchers propose different allocation rules (see, e.g., Eppen and Schrage, 1981; De Kok et al., 1994; Van der Heijden et al., 1997; Diks and De Kok, 1999) to allocate the available stock among downstream stages and to develop approximations under some form of balance assumption. Doğru et al. (2009) investigates the effect of the balance assumption with a numerical study conducted over a wide range of parameters. We also refer to Axsäter (2003) and Gallego et al. (2007) for further discussions of the issue of imbalance.

In continuous-review installation stock policies, the issue of imbalance and the stock allocation problem are not considered since one can typically assume First-Come-First-Served (FCFS) allocation rule. However, the challenge in this case is to characterise the stock-out delay, i.e. the additional waiting time of downstream stages due to eventual stock-outs at their upstream stages. The classical approach that constitutes the basis for a lot of models that consider installation stock policies is the METRIC approach of Sherbrooke (1968). He provides a simple approximate approach for repairable items (i.e. items with relatively low demand and high holding costs) controlled with continuous-review, base-stock

policies. The METRIC approach is extended to consider items with low cost and high demand and to systems that operate under different installation stock policies. There has been a lot of research in this direction; for surveys see Diks et al. (1996) and Axsäter (2003). For periodic-review installation stock policies, a classical FCFS allocation rule may not be relevant since downstream stages can place orders to their upstream echelon at the same time. Therefore, the upstream stages face the decisions of stock allocation in case of stock insufficiency. We refer the reader to Jackson (1988), Graves (1996) and Marklund and Rosling (2012) and references therein for further details on the stock allocation problem in periodic-review installation stock policies.

2.1.3 General Systems

While serial, assembly and distribution systems are extensively studied by using the SSM approach, the number of SSM formulations that consider general supply chain structures is relatively limited. This is due to the intractability of the SSM approach for general systems unless some specific assumptions and approximations. The optimal policy for such systems is unknown and at least as complex as those of distribution systems (Shi and Zhao, 2010). In what follows, we provide a comprehensive review of the different models that deal with general systems using the SSM approach.

Lee and Billington (1993) analyse a supply chain structure in Hewlett-Packard Company which constitutes a combination of assembly and distribution systems. Each stage in the considered supply chain uses a periodic-review, installation order-up-to policy. Under given service level targets for each stage, they show how to calculate the order-up-to levels by using a single-stage model. To do this, they develop approximate expressions for the replenishment time of each stage in the supply chain. The approximation is based on the assumption that in each period at most one upstream stage in the system can be out-of-stock. Under this assumption, they provide a performance analysis by evaluating various stock positioning strategies. Ettl et al. (2000) consider the same assumption for the case where each stage uses a continuous-review, installation base-stock policy, external demands follows compound Poisson processes, lead times are *i.i.d.* random variables and the system operates

with the FCFS allocation rule. They model the difference between the lead time L_j of stage $j \in \mathbf{N}$ which is a given data and its replenishment time $\tilde{L}_j$ which can be derived taking into account the probability of stock-out at upstream stages. The replenishment time $\tilde{L}_j$ of stage $j \in \mathbf{N}$ is a random variable which can be expressed as:

$$\tilde{L}_j = L_j + \max\left\{\Delta_i | i : (i, j) \in \mathbf{A}\right\}$$

where Δ_i is the stock-out delay occasioned by upstream stage *i*. The stock-out delay Δ_i is a quite intractable random variable for an exact characterisation. They derive approximations and bounds on these random variables and present an optimisation model that minimises the total inventory cost subject to customer service level constraints. This model has applied at IBM and their partners as a part of an extended-enterprise supply chain analysis tool and has yielded great benefits (see Lin et al., 2000).

Graves and Willems (2003) present an optimisation model based on the model of Ettl et al. (2000). They assume that each stage operates with a periodic-review, installation order-up-to policy with common review periods equal to one period at all stages, external demands follow stationary *i.i.d* processes and lead times are deterministic and constant. The major difference with the model proposed by Ettl et al. (2000) comes from the assumption which indicates that if an upstream stage causes a stock-out then its stock-out delay is equal to its lead time.

Simchi-Levi and Zhao (2005) also follow the approach proposed by Ettl et al. (2000) considering continuous-review, installation base-stock policies and independent Poisson processes for external demands. However, they make a different assumption on lead time of stages: while Ettl et al. (2000) consider *i.i.d.* random lead times, Simchi-Levi and Zhao (2005) consider stochastic, sequential and exogenously determined lead times. Under this assumption, orders do not cross in time. They focus on (spanning) tree network structures where there is at most one undirected path between every two stages. They develop exact recursive equations for stock-out delays at all stages of the supply chain. Guided by the exact analysis, they present approximations that improve the computational efficiency for optimisation and develop an algorithm based on dynamic programming in order to determine

the optimal or near optimal base-stock levels that minimise the total inventory cost subject to customer service level constraints. This work is extended by Zhao (2008) to more general supply chains (i.e. networks with at most one directed path between every two stages) facing compound Poisson demand. Recently, Shi and Zhao (2010) consider similar assumptions with that of Simchi-Levi and Zhao (2005) about lead times and external demand and discover some simple yet unique properties for acyclic general networks. They introduce an inventory control policy that consists of splitting inventory into multiple stock piles and dedicating each to a unique downstream path. This policy is called the dedicated stocking policy. They show that under certain assumptions, the best dedicated stocking policy always outperforms the best continuous-review, installation base-stock policy. Besides, they show that under certain conditions, an acyclic supply chain can be decomposed into a tree network structure without increasing its total inventory cost.

The work of Glasserman and Tayur (1995) is among the first which considers echelon stock policies. They assume stochastic and continuous external demands, linear backordering costs, limited production capacity at each stage and modified period-review, echelon order-up-to policies. The policy is modified in the sense that limited production capacity may preclude restoring inventories to their order-up-to levels. They show that for various cost and performance measures, derivatives with respect to echelon order-up-to levels can be consistently estimated from simulation, or even from real data. They illustrate the effectiveness of the derivative estimates by incorporating them in an optimisation procedure to find optimal order-up-to levels for a PC assembly and distribution system of a major computer manufacturer. De Kok and Fransoo (2003) introduce an echelon stock policy by extending the optimal policy described for assembly systems by Rosling (1989) to a non-optimal policy for general multi-echelon systems. They call the proposed policy the synchronised base-stock policy. Based on the insights provided for assembly systems, they translate the general system into a divergent system. Thus, they propose to apply the solution method provided by Diks and De Kok (1999) for divergent systems in order to determine the base-stock levels and rationing fractions that satisfy the customer service level constraints. This approach has been applied at Philips Electronics and has brought substantial savings that

is around \$5 million per year from the yearly turnover of \$300 million (see De Kok et al., 2005).

We have presented in this section detailed insights on models that use the SSM approach for general multi-echelon systems. The optimal policy for such systems is unknown. The available SSM literature dealing with general systems is relatively limited. The majority of papers use standard and simple installation stock policies under different assumptions regarding lead times and external demands. Besides, they usually introduce simplificative assumptions and approximations to improve the computational efficiency for the purpose of system evaluation and optimisation. This may render the model less realistic. Indeed, allowing complexity to increase may lead to better performing solutions. However, it may cause several issues not only in modelling and solution computation but also in implementation due to the complexity of operations regarding the inventory control policy to be applied.

2.2 Guaranteed-Service Model Approach

In this thesis, we have a special focus on the GSM approach since the relevant models may be computationally tractable for general multi-echelon systems and have a great potential in terms of improvement in real-world supply chains. Indeed, the research on GSM has gained interest in recent years. Although this approach was initiated more than fifty years ago with the work of Simpson (1958), almost 80% of the existing works are published in the last decade (see Figure 2.1). To the best of our knowledge, there is no literature review which gives a synthesis of the various works developed so far.

In this section, we conducted a literature review over three axes. The first axis classifies papers according to the assumptions that are considered while modelling the multi-echelon system analysed with the GSM approach (Section 2.2.1). The second axis focuses on solution techniques developed for different supply chain structures (Section 2.2.2). Finally, the third axis presents results obtained by industrial applications (Section 2.2.3). It should be noted that when an existing paper falls into more than one of these axes, the major contributions of the paper are separately presented in the concerned sections.

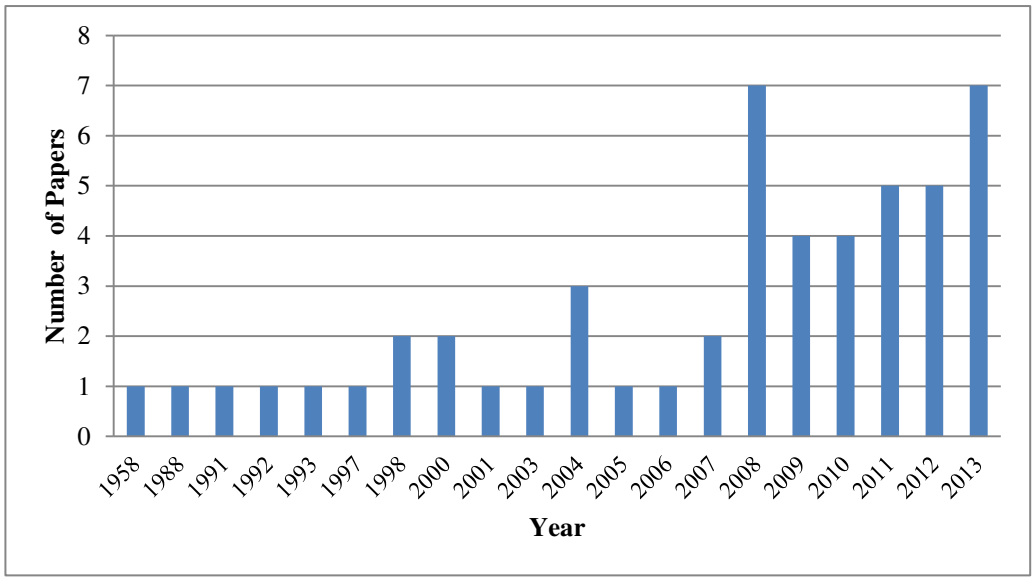

Figure 2.1: Number of publications on the GSM approach per year

2.2.1 Modelling Assumptions

This section reviews the various models that extend the original GSM to enable to capture real-world supply chain characteristics. The criteria we use to classify existing models are based on the relaxations that have been made to the original model. Indeed, these relaxations concern several assumptions used in the model and are relative to: 1) external demand, 2) lead times, 3) capacity constraints, 4) service times, 5) inventory control policies, 6) extraordinary measures and 7) decision makers. Table 2.1 gives a summary of extensions developed. The first two columns summarise the assumptions of the original GSM as presented in Section 1.2.2 of Chapter 1 while the third column lists the associated relaxations.

Original Assumptions		Associated Relaxations	Reference
External demand	Stationary demand	Non-stationary demand	Graves and Willems (2008), Neale and Willems (2009)
		Consideration of evolving demand forecasts	Schoenmeyr and Graves (2009)
	Demand bounds subject to a CSL target	Demand bounds subject to a modified fill-rate target	Inderfurth and Minner (1998)
Lead times	Known and constant	Stochastic	Inderfurth (1993), Minner (2000), Humair et al. (2013)
Capacity Constraints	None	Consideration of capacity constraints	Sitompul et al. (2008), Schoenmeyr (2008)
Service Times	Unique for all direct stages	Customer-specific service times	Graves and Willems (1998), Minner (2000)
	Constant	Dynamic	Graves and Willems (2008)
Inventory Control Policies	Periodic-review, order-up-to policy	Continuous-review batch ordering policy	Li and Chen (2012), Li et al. (2013)
	Common review periods for all stages	Arbitrary and integer	Bossert and Willems (2007)
Extraordinary measures	Non modelled explicitly	Modelling insights for these measures	Rambau and Schade (2010), Klosterhalfen and Minner (2010)
Decision-makers	Single decision-maker	Several decision-makers	Schoenmeyr (2008), Egri (2012)

Table 2.1: Classification according to modelling assumptions

External Demand

The original version of the GSM assumes that external demand comes from a stationary process and demand at each stage $j \in \mathbf{N}$ is bounded for any long period $\tau_j = 1,...,M_j$ by an increasing and concave function $D_j(\tau_j)$. Most studies adopt a Normal Distribution to represent the external demand pattern in their applications (see, e.g., Simpson, 1958; Inderfurth, 1991; Graves and Willems, 2003) and specify the demand bounds at demand stages $j \in \mathbf{N_D}$ as follows:

$$D_j(\tau_j) = \tau_j \mu_j + z_j \sigma_j \sqrt{\tau_j} \tag{2.1}$$

where z_j is the safety factor of stage *j* that relates to its CSL α_j for a cycle of length τ_j. The safety factor z_j verifies $\alpha_j = \Phi(z_j)$ where $\Phi(.)$ is the standard normal cumulative distribution function. Here, the value of z_j depends only on α_j.

For each non-demand stage $j \in \mathbf{N_S} \cup \mathbf{N_I}$, demand bounds can be set based on function (2.1). For instance, if stage *i* has a single downstream stage *j* then:

$$D_i(\tau_i) = \theta_{ij} D_j(\tau_i) \text{ for all } \tau_i = 1,..,M_i \tag{2.2}$$

If stage *i* has more than one downstream stage then the risk pooling effect that describes a relative reduction in demand variability due to combining demand of multiple downstream stages can also be incorporated into the demand bound of stage *i* (see Graves and Willems, 2000).

Inderfurth and Minner (1998) are the only ones that use another service measure than the CSL to derive the safety factor in function (2.1). They consider the modified fill-rate service measure γ'_j (see Section 1.1.5). By definition, the safety factor z_j becomes dependent on both γ'_j and τ_j. Under this setting, they show that there exists an additional coverage potential at stage *j* that represents the maximum time that can be covered without holding any safety stock at stage *j*. This can be computed based on γ'_j.

Graves and Willems (2008) assume a non-stationary demand process. They specify a planning horizon of length *H* and assume that the mean rate $\mu_j(t)$ and the standard deviation $\sigma_j(t)$ at demand stage $j \in \mathbf{N_D}$ is independent over $0 \le t \le H$. Then, they introduce the demand bound function $D_j(t - s_j^{in} - L_j, t - s_j^{out})$ defined for the time interval $(t - s_j^{in} - L_j, t - s_j^{out}]$:

$$D_j(t - s_j^{in} - L_j, t - s_j^{out}) = \int_{\tau=t-s_j^{in}-L_j}^{t-s_j^{out}} \mu_j(\tau) d\tau + z_j \sqrt{\int_{\tau=t-s_j^{in}-L_j}^{t-s_j^{out}} \sigma_j^2(\tau) d\tau} \tag{2.3}$$

Analogously to the original model, to satisfy the service time guarantee with the minimum inventory level the order-up-to level $S_j(t)$ of stage *j* at time *t* is set as follows:

$$S_j(t) = D_j(t - s_j^{in} - L_j, t - s_j^{out}) \tag{2.4}$$

Using (2.4), Graves and Willems (2008) allow the order-up-to level to vary over time and thus propose an adaptative order-up-to policy. They then replace the objective function (1.17) of the problem *P*0 (Section 1.2.2, page 30) with:

$$\min \sum_{j \in \mathbf{N}} \sum_{t=1}^{H} h_j \left(D_j(t - s_j^{in} - L_j, t - s_j^{out}) - \int_{\tau=t-s_j^{in}-L_j}^{t-s_j^{out}} \mu_j(\tau) d\tau \right) \tag{2.5}$$

Graves and Willems (2008) argue that it is reasonable to assume that (2.5) is concave.

Neale and Willems (2009) also include a non-stationary demand process in the GSM framework. They assume that the planning horizon is divided into different phases. Demand within each phase is assumed to be stationary with a known mean and standard deviation. Under this assumption, they obtain the discrete time version of demand bound function given in (2.3).

Schoenmeyr and Graves (2009) extend the non-stationary demand process assumption to evolving forecasts. Under this assumption, forecasts are regularly updated based on observed sales, advanced orders, and market intelligence. As in Graves and Willems (2008), they let the order-up-to levels vary over time and they call this replenishment policy as "forecast-based ordering policy". They demonstrate that under the considered setting, the concavity property is still valid.

Regarding the external demand assumption, researchers have so far proposed extensions by considering non-stationary and evolving demand patterns. As will be further presented in Section 2.2.3, these extensions enable to realise real-world applications since many companies experience non-stationary demand because of short product life cycles, seasonality, sales-force incentives etc.

Lead Times

A second important extension concerns the relaxation of the deterministic lead time assumption. Different approaches exist in the literature in order to incorporate stochastic lead times into the GSM framework.

Inderfurth (1993) develops the first approach by considering a general multi-echelon system. He assumes that the lead time at every stage $j \in \mathbf{N}$ is normally distributed with mean L_j and standard deviation σ_{L_j}. He proposes to apply a well-known result used in single-echelon theory (see e.g. Silver et al., 1998) to calculate the safety stock at demand stage $j \in \mathbf{N_D}$:

$$z_j \sqrt{L_j \sigma_j^2 + \mu_j^2 \sigma_{L_j}^2} \tag{2.6}$$

At non-demand stages, he classifies the lead time variability into reasonable and extraordinary variations. He then specifies the planned lead time at non-demand stage $j \in \mathbf{N_S} \cup \mathbf{N_I}$ with:

$$L'_j = L_j + z_{L_j}\sigma_{L_j}$$

where z_{L_j} relates to the service level which denotes the probability that the lead time realisation does not exceed the planned lead time L'_j. The original problem $P0$ is then modified by replacing L_j with L'_j for non-demand stages in (1.17) and in (1.18). In this setting, the objective function becomes:

$$\min \sum_{j\in\mathbf{N_S}\cup\mathbf{N_I}} h_j z_j \sqrt{(s_j^{in} + L'_j - s_j^{out})\sigma_j^2} + \sum_{j\in\mathbf{N_D}} h_j z_j \sqrt{(s_j^{in} + L_j - s_j^{out})\sigma_j^2 + \mu_j^2\sigma_{L_j}}$$

Minner (2000) shows that the approach of Inderfurth (1993) leads to large safety stocks. Besides, it requires the specification of two different service levels for both demand and lead time. He thus proposes another approach where only one service level has to be chosen to dimension safety stocks. He considers a serial system where stages are numbered from 1 to n from the most upstream to the most downstream stage. The cumulative net replenishment times from stage 1 to stage j are restricted with the sum of average lead times plus a safety surplus. In case of independent lead times, the safety surplus of stage $j \in \mathbf{N}$ is given by:

$$\left(\frac{\mu_j}{\sigma_j}\right)^2 \sum_{i=1}^{j} \sigma_{L_i}^2$$

Using this approach, the objective function (1.17) remains the same, whereas safety surplus parameters appear in the constraints. The objective function remains concave using models proposed by Inderfurth (1993) and Minner (2000).

Recently, Humair et al. (2013) contributes to the literature with a more sophisticated approach to deal with stochastic lead times in general multi-echelon systems. They define the shortfall at stage j as the difference between what stage j has shipped out and what it has replenished. Because demands and lead times are random, shortfalls and net replenishment

times are also random variables. Shortfall of stage j is positive only if the realised net replenishment time is positive. They assume that the positive shortfall has an approximately normal distribution. Hence, they set the safety stock at stage j as:

$$z_j \sqrt{Q((s_j^{out} - s_j^{in})^+)\sigma_j^2 + \mu_j^2 R((s_j^{out} - s_j^{in})^+)} \tag{2.7}$$

where $(s_j^{out} - s_j^{in})^+ = \max\{s_j^{out} - s_j^{in}, 0\}$ and functions $Q(.)$ and $R(.)$ are the mean and variance of the net replenishment time random variable conditional on being positive. We can observe that the safety stock expression in (2.6) and (2.7) are structurally similar. The only difference is that lead time parameters are replaced with equivalent functions regarding the positive part of the net replenishment time random variable. Furthermore, Humair et al. (2013) show that when a stage's service time exceeds its inbound service time, some lead time realisations might cause orders to arrive at the relevant stage before the associated downstream demand has shipped. This forces some stages to carry an additional stock that is called the early arrival stock. Humair et al. (2013) define the objective function as the total cost of safety stock and early arrival stock. This setting might cause that the objective function becomes non-concave or non-differentiable.

Concerning the different approaches proposed to incorporate stochastic lead times into the GSM model, we notice that a straightforward modelling approach may lead to significant safety stocks whereas a more sophisticated approach may violate the concavity property of the objective function and hence increase the computational complexity of the problem. Indeed, when implementing these approaches, the objective is to reach the best trade-off between the quality of the solution and the computational complexity.

Capacity Constraints

In the original GSM, there are no capacity constraints that limit the quantity of items which can be processed at stages. To the best of our knowledge, Sitompul et al. (2008) is among the first who take the capacity constraints into account. They define q_j as the maximum quantity of items which can be processed at stage j during one period. They consider the CSL as the service measure and Normal Distribution as demand pattern. They

show that the safety stock level needs to be increased by a correction factor φ_j which depends on the excess capacity $(q_j - \mu_j)$ over the standard deviation of demand during the net replenishment time. Hence, instead of (2.1), they set the demand bound $D'_j(\tau_j)$ at demand stage *j* as follows:

$$D'_j(\tau_j) = \tau_j \mu_j + \varphi_j z_j \sigma_j \sqrt{\tau_j}$$

They use exponential interpolation to estimate the correction factor φ_j and test their approach with Monte Carlo simulation. However, their work remains approximate and their findings are based on a limited set of simulations.

Schoenmeyr (2008) considers independently the same relaxation and provides an exact derivation of demand bounds. He demonstrates that under capacity constraints demand bound functions $D'_j(\tau_j)$ can be defined as:

$$D'_j(\tau_j) = \max_{m \in \mathbf{Z}} \{D_j(\tau_j + m) - q_j m\} \tag{2.8}$$

where *m* belongs to the set of non-negative integers $\mathbf{Z}$. In (2.8), function $D_j(\tau_j + m)$ can be defined as in (2.1) or by using other approaches presented in the "External Demand" subsection. Under the order-up-to policy assumption, Schoenmeyr (2008) proposes to replace demand bound functions used in (1.17) by (2.8). He proves that the concavity property of the objective function is still valid for this extension.

Schoenmeyr (2008) also suggests an appropriate inventory control policy that is called "censored order policy" where a stage does not place a full order to the upstream stage if it knows that it will be unable to process such a quantity because of its capacity constraint. Therefore, the order-up-to level is set as:

$$S_j(\tau_j) = \min\{D'_j(\tau_j), q_j \tau_j\}$$

Numerical experiments show that the proposed policy is better than the order-up-to policy in terms of cost. However, the censored order policy gives better results even in the absence of capacity constraints. This paradox can be explained by the smoothing effect of this policy which reduces demand variability at upstream stages. This shows that the order-up-to policy is not necessarily optimal for a multi-echelon system with guaranteed service.

Service Times

Each stage offers the same service time for its customers in the original GSM. Graves and Willems (1998) describe how to transform the model to permit customer-specific service times. The main idea is to insert dummy stages (nodes) between downstream-upstream stages, so that each stage still quotes the same service time to its downstream stages. Dummy stages have zero cost and zero lead time but they are free to quote any valid service time. In this approach, the original optimisation model structure does not change whereas the considered network is modified inserting dummy nodes. Minner (2000) considers the same relaxation for distribution systems and proposes two different model formulations. In the first model, he provides a natural extension by allowing stage *i* to quote different service times s_{ij}^{out} for all of its direct downstream stages $j:(i,j)\in \mathbf{A}$. In the second model, stage *i* quotes a service time s_{ij}^{out} which is differentiated according to the requests of the demand stage $j \in \mathbf{N_D}$ if there is a path between *i* and *j* in the network. He then compares these two models and shows that the second one dominates the first one in terms of cost performance. However, in terms of computational effort and ease of implementation the first model is preferred to the second one.

In the original model, service times are assumed to be constant over time. Neale and Willems (2009) justify this assumption since they lead to constant safety stock locations, simplify the model and reduce computational requirements. However, Graves and Willems (2008) show that constant service times may lead to sub-optimal solutions when demand is non-stationary. Their study is based on a chosen example and does not give a general idea about the sub-optimality of constant service time solutions.

Inventory Control Policies

As presented in the previous subsections, researchers propose different replenishment policies for situations that involve non-stationary demand and capacity constraints. In this section, we focus on works in which the major contribution is the relaxation of the original replenishment policy assumption of the GSM.

The original GSM assumes a periodic-review, constant order-up-to level policy with a common review period for all stages. Bossert and Willems (2007) extend the GSM to allow

review periods to be stage-dependent. They assume that the review period of each stage has an arbitrary integer and known value. A simple approach to integrate this assumption would be to aggregate the review periods into lead times. However, they propose better replenishment policies in terms of cost by examining the cyclic inventory dynamics. Policies that are proposed include constant and adaptative order-up-to and constant safety stock targets.

Li and Chen (2012) consider a continuous-review policy in the GSM framework. They study a serial system with Poisson external demand. They assume that each stage operates with an echelon order point, order quantity policy. Li et al. (2013) extend this work to assembly systems. Since they attempt to optimise both order point and order quantity parameters, the resulting objective function is neither convex nor concave. They decompose the relevant problem into two sub-problems, the order quantity decision sub-problem and the order point decision sub-problem. In their setting, the two sub-problems are independent where the order point decision sub-problem is equivalent to the original GSM.

Schoenmeyr (2008) shows that the periodic-review, order-up-to policy is not necessarily optimal for multi-echelon systems modelled with the GSM framework. In his work, he gives some insights about the sub-optimality of this assumption. He shows that a new replenishment policy that sets upper limits for the ordered quantity may lead to a lower total inventory cost. Such a policy at demand stages smoothes orders placed to the upstream stages and absorbs demand variability. However, he considers the upper limits for ordered quantities as given parameters and does not discuss how to specify them in practice if there are no evident limits such as capacity constraints in the system.

Extraordinary Measures

The original GSM does not explicitly model what happens when demand exceeds the demand bounds. The implicit assumption in most of the GSM extensions is that only demand within the specified demand bounds is propagated through the system whereas demand in excess of these bounds is truncated and handled outside the normal supply chain.

Only few results are available regarding the impact of extraordinary measures on the considered systems. Klosterhalfen and Minner (2010) assume that the unbounded original external demand is propagated through the system. They model the use of express expediting as extraordinary measure, i.e. accelerated transportation by speeding up the excess amount from the stage's own pipeline stock. They consider a simplified assumption regarding the specification of the cost associated with express expediting. That is, each unit of item expedited in express incurs a cost irrespective of the delay to meet the associated demand. Their simulation results show the relevancy of this assumption. Besides, they show that the use of express expediting decreases the pipeline stock and they include the associated reduction in the cost function. Under the proposed setting, they provide an extension with concave objective function for given safety factors. Rambau and Schade (2010) include extraordinary measure costs for both delays and unmet demand and propose a stochastic programming version of the GSM. However, they neither specify what measure among express expediting, overtime production, subcontracting etc. is considered nor model the eventual impact of these measures on inventory dynamics.

Decision-Makers

In the original GSM, it is implicitly assumed that there exists a single decision-maker that takes the safety stock placement decision for the whole supply chain. Schoenmeyr (2008) argues that the GSM framework is also relevant when different parts of the supply chain are controlled by different decision-makers which may have competing and conflicting interests. He proposes a simple contract structure in order to facilitate the relationship between two decision-makers that control their part according to the GSM approach. Under the proposed contract, downstream holding costs are not affected by the non-value added mark-up applied by the upstream decision-maker. The proposed contract is incentive compatible, i.e. once the decision-makers have agreed on the global optimal service time, it will be in their own best interests to operate the supply chain according to the globally optimal solution.

Similarly, Egri (2012) considers the GSM approach in a decentralised serial supply chain involving a number of autonomous stages where each stage represents a different

decision-maker. He proposes different mechanisms that facilitate the alignment of conflicting goals and achieve the global optimal solution in such systems. Numerical examples provided illustrate how the proposed mechanisms enable to achieve the globally optimal solution in different ways.

2.2.2 Solution Methods Developed

This section aims at classifying the existing literature according to the solution methods developed. Table 2.2 gives a summary of the existing works regarding the nature of their contribution (column 1), the supply chain network structure considered (column 2), the objective function (column 3) and the method used (column 4).

Simpson (1958) considers a serial network and proves that optimal solutions of the problem $P0$ can only occur on the extreme points of the solution set since this problem is the minimisation of a concave function over a closed, bounded convex set. This property is called the all-or-nothing property or the extreme point property. In solutions holding this property in a serial system, a stage has either no safety stock or has sufficient safety stock to decouple it from its downstream stage. Inderfurth (1991) shows the validity of this property for distribution networks. Inderfurth and Minner (1998) present the optimal solution properties for serial, assembly and distribution networks under different service measures. Lesnaia (2004) shows the optimality conditions for spanning tree networks in which there is at most one path of arcs between two nodes. She also extends these results to provide the optimality conditions for a special case of general networks called networks with Clusters of Commonality (CoC). In such networks, when each cluster is replaced by a single node, the resulting network is a spanning tree. Minner (2000) characterises the optimal solution properties for general acyclic networks. Finally, Minner (2001) discusses the optimal solution properties for general cyclic networks and presents an extended extreme point representation with the synchronisation of service times and item returns in such systems.

Considering the optimal solution properties for different supply chain networks, optimal solutions of the problem $P0$ can be determined by enumeration. However, the extreme point property enables the development of more efficient exact solution methods.

For instance, Inderfurth (1991) introduces the dynamic programming algorithm for distribution systems based on this property (see also Inderfurth, 1992). Similarly, Minner (1997) presents dynamic programming algorithms for serial, assembly (*Ass.*) and distribution (*Distr.*) systems under different service measures. The dynamic programming approach is extended by Graves and Willems (2000) to consider spanning tree networks. Indeed, these networks represent the most complex type of multi-echelon system to which the single-state variable dynamic programming approach can be directly applied. The computational complexity of the algorithm proposed by Graves and Willems (2000) is of order nM^2 where n is the number of nodes and M is the maximum replenishment time in the system. Lesnaia (2004) improves this algorithm by considering the relevant optimal solution properties and proposes a dynamic programming algorithm of order n^3 for such networks. Besides, she also develops a branch and bound algorithm for networks with CoC using the optimal solution properties for these networks. Minner (2000) shows that general acyclic networks lead to higher dimensional states and decision spaces to use the dynamic programming algorithm. He gives insights on both forward recursion and backward recursion dynamic programming algorithms for these systems. To minimise the number of state variables, he argues that a backward recursion is preferable if the supply chain network represents distribution dominance whereas a forward recursion is advantageous in case of assembly dominance. Lesnaia (2004) shows that the general acyclic network problem is NP-hard. She develops a branch-and-bound algorithm to solve the problem to optimality. Magnanti et al. (2006) provide an exact solution method for this problem without referring to the optimal solution properties. They use successive piecewise linear approximation to obtain tight approximation to the concave objective function. Hence, the problem is transformed to a Mixed Integer Programming problem and is solved by a commercial solver.

Researchers have also proposed approximate (*Appr.*) solution approaches in order to find near-optimal solutions with less computational effort for general acyclic network problem with concave objective function. Minner (2000) investigates several heuristic approaches such as Linear Approximation, Simulated Annealing, Threshold Accepting and Tabu Search. Shu and Karimi (2009) propose heuristic approaches following the idea of

Magnanti et al. (2006). They suggest two efficient heuristics: the first one uses continuous approximation while the second one employs a two-piece linear approximation to approximate the concave objective function. For large and dense acyclic networks, these heuristic algorithms are roughly faster than the exact method proposed by Magnanti et al. (2006) and give solutions within 7% and 4% of the optimum on average. Besides, the performance of their methods in terms of solution quality is nearly independent from the network size. In the same setting, Li and Jiang (2012) propose a heuristic approach integrating constraint programming with a genetic algorithm. They compare their solution method with heuristics of Shu and Karimi (2009). Their method offers a novel solution approach that balances solution speed and quality while heuristics of Shu and Karimi (2009) are better in terms of computation efficiency. Besides, they show that the quality of heuristics of Shu and Karimi (2009) varies significantly with the maximum service times of demand stages. However, their solution method appears robust with respect to these parameters.

Solution methods that we have presented so far assume a concave form for the objective function (1.17). These methods can also be used to solve the extensions presented in Section 2.2.1 if the relevant objective function is assumed to be concave (e.g., Inderfurth, 1993; Graves and Willems, 2008; Neale and Willems, 2009). However, they cannot be used when the model includes, e.g., fixed costs associated with holding inventory, non-nested review periods or stochastic lead times (as presented in Humair et al., 2013). To study such cases, arbitrary cost functions for which there are not any structural limitations such as concavity or monotonicity are considered by Humair and Willems (2006). They focus on networks with CoC and present an exact solution approach based on dynamic programming. Similarly, Humair and Willems (2011) consider arbitrary stage cost functions for general acyclic network problem. They provide an exact solution approach extending the dynamic programming algorithm developed by Graves and Willems (2000). They also present two significantly faster and near-optimal heuristics. They test their approaches on a published data set of real-world supply chains (Willems, 2008) in order to demonstrate the performance and the consistency of their algorithms.

We notice that solution methods for general acyclic networks have been largely studied in the literature. However, a complete performance evaluation for these methods has not been provided yet. Besides, general cyclic networks have attracted much less attention.

Reference	Nature of the Contribution			Supply Chain Network							Objective Function		Method used
	Optimal Solution Property	*Exact Solution Approach*	*Appr. Solution Approach*	*Serial*	*Ass.*	*Distr.*	*Spanning Tree*	*Networks with CoC*	*General Acyclic*	*General Cyclic*	*Concave*	*Arbitrary*	
Simpson (1958)	x			x							x		Extreme point property
Inderfurth (1991)	x	x				x					x		Dynamic programming
Minner (1997)	x	x		x	x	x					x		Dynamic programming under different service measures
Inderfurth and Minner (1998)	x			x	x	x					x		Optimal solution properties under different service measures
Graves and Willems (2000)		x					x				x		Dynamic programming
Minner (2000)	x	x	x	x	x	x			x	x	x		Dynamic programming, linear approximation, simulated annealing etc.

Table 2.2: Classification according to contributions regarding the optimal solution properties and solution methods

Reference	Nature of the Contribution			Supply Chain Network							Objective Function		Method used
	Optimal Solution Property	*Exact Solution Approach*	*Appr. Solution Approach*	*Serial*	*Ass.*	*Distr.*	*Spanning Tree*	*Networks with CoC*	*General Acyclic*	*General Cyclic*	*Concave*	*Arbitrary*	
Minner (2001)	x									x	x		Extreme point property
Lesnaia (2004)	x	x					x	x	x		x		Branch-and-bound algorithms
Humair and Willems (2006)		x						x				x	Dynamic programming
Magnanti (2006)		x							x		x		Successive piecewise linear approximation
Shu and Karimi (2009)			x						x		x		Linear and two-piece linear approximations
Humair and Willems (2011)		x	x						x			x	Dynamic programming
Li and Jiang (2012)			x						x		x		A hybrid algorithm integrating constraint programming and genetic algorithm

Table 2.2: Classification according to contributions regarding the optimal solution properties and solution methods (continued)

2.2.3 Industrial Applications

In this section, we focus on papers that apply the GSM approach to real industrial settings. In Table 2.3, we present for different industries/companies (columns 1 and 2) results obtained (column 5) by the GSM based models. Besides, we distinguish two types of papers (column 4). A first category of papers (Type I papers) deals with the practical application of a GSM extension in a specific company. Among examples of companies for which successful applications are realised, one can cite Eastman Kodak, Hewlett-Packard (HP), Intel and Procter & Gamble (P&G). A second category of papers (Type II papers) solves an integrated GSM where the multi-echelon inventory optimisation problem is jointly considered with another supply chain decision such as the supply chain configuration or production planning problems.

As such, among examples of papers that lie in the first category, Billington et al. (2004) show that savings realised by using the GSM approach for Hewlett-Packard's Digital Camera and Inkjet Supplies business exceed $130 million. Farasyn et al. (2011) also report that multi-echelon models based on the GSM approach now drive 30% of Procter & Gamble's business and have produced 7% of average inventory reduction. Wieland et al. (2012) describe a multi-echelon inventory optimisation project at Intel and indicate that after its implementation, inventory levels are reduced more than 11% providing average service levels exceeding 90%.

In the second category of papers (Type II papers), a first example of decisions that are jointly studied with the inventory optimisation problem modelled under a GSM approach is the supply chain configuration problem. More precisely, this problem consists of determining, for a given stage, which option to select among different alternatives (e.g., which supplier to choose, which transportation modes to use etc.) where each alternative differs in cost and lead time that in turn, impact safety stock related decisions. Graves and Willems (2003) is among the first who integrates the GSM approach into the supply chain configuration problem. Hence, the proposed model aims at simultaneously determining the best options to select as well as the related safety stock placements in the supply chain. The global cost formulation considered includes safety stock, pipeline stock and processing costs.

They show on a real bulldozer supply chain that by reconfiguring the supply chain, the total supply chain cost may be decreased compared to the original GSM solution with predefined standard options. Graves and Willems (2005) present a dynamic programming algorithm to solve this problem for spanning tree networks. They apply this model to a notebook computer supply chain and show that by optimising the supply chain configuration, the company saves more than $2.2 million compared to the current policy. Li and Womer (2008) consider also a similar problem assuming that each option may also be differentiated by the level of quality provided and resource capacities consumed. Hence, they add to model constraints a certain threshold of quality level that has to be maintained for the entire supply chain and resource capacities regarding materials, budgets, machines, vehicles, personnel etc. They model the problem as a project scheduling problem considering the objective function of minimising the cycle time of a new final product, i.e. the time required to manufacture and distribute the new product to customers. They propose a constraint programming based solution approach to solve the problem for general networks. You and Grossmann (2008; 2010) consider a tri-echelon distribution network and add to the model of Graves and Willems (2005) fixed installation costs for certain stages (called distribution centres). They develop decomposition algorithms that obtain optimal or near-optimal solutions. You and Grossmann (2008) present illustrative examples in industrial chemicals industry to show the trade-offs in this problem by considering different parameters for transportation and inventory costs. For a similar problem, Nepal et al. (2011) provide a multi-objective optimisation model by considering the minimisation of the total supply chain cost and the maximisation of the total compatibility index for the selected options (i.e., selected members, firms, partners when configuring the supply chain). The compatibility index of different options is a subjective parameter in nature and is assumed to be determined regarding the structural, managerial and financial aspects of the associated members. Nepal et al. (2011) formulate this problem as a weighted goal programming model and propose a genetic algorithm to obtain near-optimal solutions. Funaki (2012) extends the work of Graves and Willems (2005) to consider non-stationary demand and to include the due-date requirements of customers. They provide an approximate optimisation approach to solve this problem and demonstrate its effectiveness on real-world

examples. Another related work to that of Graves and Willems (2005) is provided by Klosterhalfen (2010). For companies relying on two supply options, he integrates dual-sourcing into the GSM framework. Instead of choosing only a single supply option, they assume that a certain fraction of demand may be allocated to each supplier in every period (order-splitting policy). Their findings confirm nevertheless that choosing a single option as in Graves and Willems (2005) is often reasonable since the cost advantage of dual sourcing is not very significant for the considered problems.

A second decision analysed simultaneously with a GSM based multi-echelon inventory optimisation problem is the production planning problem. Tian et al. (2011) propose an iterative approach to jointly solve the problem of allocating production capacity and determining safety stock levels at different stages. They prove the applicability of their solution method on industrial-scale problems through real-world examples in the semiconductor industry. You and Grossmann (2011) propose another formulation for a similar problem in the chemical process industry. The proposed model simultaneously determines the optimal purchase amount of each raw material, production levels in each process, sale amount of each final product, internal demand of each production process and safety stock level of each chemical in the considered network. To solve efficiently this problem for large and complex supply chains, they exploit some model properties and propose an exact solution method.

Papers presented in this section show that the deployment of the GSM approach in industry yields great benefits. Nowadays, the company Logility offers a multi-echelon inventory optimisation software tool that enables companies to implement some of these models. We assume that their tool most probably employs the GSM approach since the affiliated scientist for this tool contributes to the GSM literature. Indeed, the benefit from the implementation of multi-echelon inventory optimisation tools in companies is estimated as 3.1% service level improvement and 15% decrease in cash-to-cash cycle (Aberdeen Group, 2012). Results that we present in this section confirm this estimation for the GSM approach based implementations (see also Table 2.3).

Industry	Company	Reference	Paper Type	Main Results
Computer Hard-ware	Hewlett-Packard	Billington et al. (2004)	Type I	Total supply chain cost reduction of over $130 million while maintaining high service levels
	Not specified	Graves and Willems (2005)	Type II (Supply chain configuration)	Saving of $2.2 million compared to the solution obtained by the current policy
	Not specified	Li and Womer (2008)	Type II (Supply chain configuration)	Same application with Graves and Willems (2005), managerial insights on understanding the benefit of increasing resource capacity and system reliability
	Microsoft	Neal and Willems (2009)	Type I	18%-20% increase in inventory turns, 6%-7% increase in fill rates
Consumer Goods	Procter & Gamble	Farasyn et al. (2011)	Type I	7% inventory reduction on average for 30% of their business
	Not specified	Humair et al. (2013)	Type I	Simplified assumptions under stochastic lead times may overestimate (2.4%) or underestimate (-3.2%) total inventory required
Digital Imaging	Eastman Kodak	Graves and Willems (2000)	Type I	Total inventory cost reduction of over one third while increasing service levels
Industrial Chemicals	Celanese	Bossert and Willems (2007)	Type I	30% decrease in inventory levels compared to a simple modelling approach in case of stage dependent review periods
	Not specified	You and Grossmann (2008)	Type II (Supply chain configuration)	Illustrative examples to show the trade-off results under different cost parameters
	Not specified	You and Grossmann (2011)	Type II (Production Planning)	Case studies to demonstrate the performance of the proposed exact solution method in terms of computational time
	Not specified	Humair et al. (2013)	Type I	Using simplified assumptions under stochastic lead times may overestimate (4%) or underestimate (-18%) the total inventory cost
Machinery	Not specified	Graves and Willems (2003)	Type II (Supply chain configuration)	Total supply chain cost decrease of 0.38% compared to the original solution
	Case New Holland	Neal and Willems (2009)	Type I	Total supply chain inventory reduction of over 20%
	Not specified	Funaki (2010)	Type II (Supply chain configuration)	Examples to illustrate significance of the proposed model for practical situations and effectiveness of the proposed solution approach
Semiconductor	Not specified	Tian et al. (2011)	Type II (Production Planning)	Illustration of the proposed optimisation approach and its applicability
	Intel	Wieland et al. (2012)	Type I	11% decrease in inventory levels while providing service levels exceeding 90%
Electronic Test Equipment	Theradyne	Schoenmeyr and Graves (2009)	Type I	Total safety stock cost reduction of 25% by incorporating the forecasting process into the GSM

Table 2.3. Classification according to industrial applications

2.3 Guaranteed-Service vs. Stochastic-Service Approach

In the literature, only few contributions are provided concerning the comparison of the SSM and GSM approaches in terms of cost. Graves and Willems (2003) illustrate the contrast between the two approaches on real-world applications. Klosterhalfen and Minner (2007) provide a comparison of the two approaches for two-stage serial systems incorporating the cost of extraordinary measures into the GSM. They then extend this work to two-echelon distribution systems (Klosterhalfen and Minner, 2010). Their simulation results show that the cost difference between the two approaches is not very large (4% at most) and the GSM approach have a better performance for moderate cost of extraordinary measures, large lead times at the upstream echelon and high service level targets at the downstream echelon. Recently, Klosterhalfen et al. (2013) develop a hybrid-service approach that combines the SSM and GSM approaches. For each stage, they determine the best approach between the SSM and the GSM to minimise the total inventory cost. The proposed hybrid-service approach not only mitigates the risk of choosing the wrong approach but also improves solutions that would be obtained using only one of the two approaches.

Some practical differences are observed in implementation of the two approaches in industry. In the SSM approach, backorders may cause variability of deliveries. According to Minner (2000), even the inventory control policy for overall system is optimised, due to the variability of deliveries at a certain stage, a local manager might feel the need of more safety stocks and might deviate from the optimal solution. In the GSM approach, this shortcoming is avoided by the guaranteed-service time assumption. According to Graves and Willems (2000) managers seem more comfortable with the notion of guaranteed-service time.

2.4 Conclusion

In this chapter, we reviewed various works pertaining to the SSM and GSM approaches. First, we classified the works pertaining to the SSM approach according to the system structure considered since the models and results differ greatly according to this aspect. For serial and assembly systems, exact results can be obtained. Some key assumptions are required to find exact results for distribution systems. For general systems, the structure

of optimal policies is unknown and characterising the optimal policy is at least as complex as for those of distribution systems. There exist only few papers that deal with general systems using the SSM approach. The majority of these papers consider standard and simple installation stock policies under different assumptions regarding lead times and external demand.

Second, we conducted a comprehensive review of the GSM literature by proposing a classification along three axes. The first axis is dedicated to models that extend the original model by relaxing some of its oversimplified assumptions. The second axis presents solution techniques that have evolved in order to be applicable for large and complex systems. The third axis presents industrial applications of the GSM as well as the benefits they have yielded.

For the purpose of computational efficiency in general multi-echelon structures, the SSM and GSM approaches both introduce several assumptions. However, the nature of assumptions used in the GSM is different to those of the SSM approach. The SSM usually makes assumptions to approximate the replenishment times, i.e. stock-out delays of stages whereas the GSM uses the bounded demand assumption that enables the introduction of guaranteed-service time notion and hence renders the stock-out delays equal to zero. A subtle point regarding the GSM is the specification of demand bounds. This should be made considering several factors such as the behaviour of customers in stock-out situations, the availability and the cost of extraordinary measures at different stages. Nevertheless, the bounded demand assumption is crucial to obtain a computationally tractable model for general multi-echelon systems. Since the GSM enables to deal with large and complex structures, it is more frequently applied in real-world supply chain settings. In contrast, the SSM research mostly focuses on serial, assembly or two-echelon distribution systems and its deployment in industry is relatively limited.

In this thesis, we aim at providing contributions to the multi-echelon inventory optimisation problem using the GSM approach since this approach has a great potential in terms of improvement in real-world supply chains. Our review allows us identifying some gaps in the GSM literature. First of all, we noticed that the impact of some simplificative

assumptions of the GSM is not elaborately studied. In particular, a comprehensive analysis is required to understand the cost and service level impact of certain assumptions regarding demand bounds, guaranteed-service times and review periods. We provide such an analysis in Chapter 3. Second, this review shows that existing models consider review periods as given input parameters of the problem. We present in Chapter 4 how to incorporate ordering costs into the GSM in order to optimise the safety stock levels (order-up-to levels) and review periods (reorder intervals) simultaneously. Third, this review reveals that demand bound of each stage is determined based on a safety factor value. However, it is not obvious how to specify the safety factors at different stages so that the target service levels are met at the most downstream stages. This issue becomes even more important in decentralised supply chains. Chapter 5 presents this problem.

CHAPTER 3: QUANTIFYING THE IMPACTS OF THE GUARANTEED-SERVICE MODEL ASSUMPTIONS

The development of computationally tractable approaches for optimising the placement and the amount of safety stocks in multi-echelon inventory systems is a complex task. The GSM has some specific assumptions that render the model deterministic and that enable the consideration of complex and large multi-echelon systems. Section 2.2.1 of Chapter 2 reviews models that attempt to relax some of these assumptions. Among these works, we can find models that consider non-stationary external demands, stochastic lead times, capacity constraints, different inventory control policies etc. Their main objective is to make the GSM more realistic and to capture real-world supply chain characteristics. After all, the literature still lacks some studies that aim at quantifying the consequences associated with several underlying assumptions of the GSM such as bounded demand, guaranteed-service times and common review periods assumptions. In this chapter, we investigate the impact of these assumptions on customer service levels or safety stock costs for serial and assembly systems.

The analysis is conducted individually for each assumption; in each section we present the results associated with a single assumption. First, we discuss the bounded demand assumption and its impact on customer service levels. In the GSM, the demand bound at a stage represents the maximum amount of demand that can be satisfied from the stock of this stage during its net replenishment time. In the literature, demand bounds are usually specified based on a target customer service level. Our analysis shows that the effectively observed service level at a demand stage would usually be less than the target one under the GSM setting. The gap may be significant in real-world systems.

Second, we discuss the impact of the guaranteed-service time assumption. This assumption implies that each stage quotes a guaranteed-service time to its customers and provides 100% service for these service times. Hence, there are no backorders between customer-supplier stages. Our analysis shows that the total safety stock cost obtained under this assumption may be significantly higher than the one obtained without this assumption.

Third, we examine the impact of the common review periods assumption. In practice, review periods can be stage-dependent, i.e. they can differ from stage to stage considering the economies of scale and/or the availability of resources. We show how to incorporate stage-dependent nested review periods into the GSM by developing an appropriate expression for the demand bounds of stages. The analysis conducted shows that the original GSM may cause a significant safety stock cost increase for long review periods.

For the numerical analysis of each assumption, we use the same test problems. First, a five-stage serial system is considered. The impact of each assumption is tested under different combinations of lead time, stage cost and target customer service level alternatives. These test problems are inspired from Schoenmeyr (2008). Second, a real-world assembly system previously presented by Graves and Willems (2003) is considered. Using this test problem, we illustrate the consequences of the GSM assumptions in a real-world problem setting.

This chapter is divided into 4 sections. Section 3.1, Section 3.2 and Section 3.3 present the analysis conducted for the bounded demand, guaranteed-service times and common review periods assumptions, respectively. Section 3.4 draws some conclusions and suggests future research directions.

3.1 Bounded Demand Assumption

In the GSM, one of the key assumptions indicates that demand is bounded at each stage of the supply chain. Indeed, the bounded demand assumption does not imply that demand can never exceed the specified demand bounds. The GSM model deals with the demand uncertainty by dividing demand variations into two ranges, ordinary and extraordinary demand variations. Safety stocks are dimensioned to deal with ordinary demand variations, i.e. to cover the demand part that stays within the specified demand bounds. Extraordinary demand variations occur when demand exceeds the specified demand bounds. Most of existing works does not address what happens in case of extraordinary demand variations. These works implicitly assume that only demand within the specified demand bounds is propagated through the system whereas demand in excess of these bounds is truncated and handled outside the normal supply chain. In this section, we consider this

common assumption and examine its impact on customer service levels. In particular, we investigate the service level that results from holding safety stocks across the supply chain under this assumption. That is, we assess the effectively observed service level at demand stages (most downstream stages that faces the external demand) when demand is truncated according to the specified demand bounds.

In the GSM literature, most studies specify the demand bounds in terms of a target non-stock-out probability, i.e. based on a target Cycle-Service-Level (CSL). The effectively observed CSL in the considered systems can be defined as the probability that the safety stocks in the system cover demand variations. Under the existence of demand bounds, the effectively observed CSL at demand stages may be less than the one used to define the demand bounds. The gap is due to the fact that the CSL at a demand stage is affected by the demand bounds applied at its upstream stages. Particularly, this may happen when the net replenishment times of upstream-downstream stages are different. This issue is not elaborately studied in the GSM literature.

In what follows, first, we formalise the satisfied demand function and the effectively observed CSL (Section 3.1.1). Second, we illustrate the CSL deviation on a simple example and provide an approximation (Section 3.1.2). Third, we present the test problems used in numerical analyses along this chapter (Section 3.1.3). Forth, we carry out a simulation study in order to assess the service level deviation for the considered test problems (Section 3.1.3).

3.1.1 Effectively Observed Cycle-Service-Level

In this section, we formulate the effectively observed CSL and the satisfied demand function under the commonly used GSM assumption which indicates that the part of the arrival demand in excess of the specified demand bounds is truncated and handled outside the normal supply chain. We ignore the effect of extraordinary measures on customer service level, i.e. we focus on the service level that results from holding safety stocks across the supply chain. First, we provide the formulation of the effectively observed CSL for serial systems then; we show how to extend it to more general systems. Besides, by using the

results provided in this section, we present a remark on the derivation of expected inventory levels in the Appendix of this chapter.

Let consider an *n*-stage serial system where stages are increasingly numbered from upstream to downstream. Without loss of generality, the coefficient of utilisation $\theta_{j-1,j} = 1$ for each $j \in \{2,...,n\}$. The demand bounds at stages $j \in \{1,2,..,n\}$ are specified by function $D(.)$ based on a target CSL α. Since we have a single demand stage for the considered system, we can omit stage indexes of demand bounds for ease of exposition. Hence, the use of (2.1) and (2.2) (Chapter 2, pages 41-42) in the considered system leads to the following:

$$D(\tau_j) = \mu\tau_j + z\sigma\sqrt{\tau_j} \text{ for } j \in \{1,2,...,n\} \quad (3.1)$$

The optimal net replenishment time τ_j^* of each stage $j \in \{1,2,..,n\}$ is obtained by solving the problem *P*0 formulated in Section 1.2.2 of Chapter 1 (page 30). The optimal order-up-to levels S_j^* correspond to the existing demand bounds $D(\tau_j^*)$ over τ_j^* consecutive periods since we have $S_j^* = D(\tau_j^*)$. In practice, it is not necessary to impose a demand bound over $\tau > 0$ consecutive periods if $\tau \neq \tau_j^*$ for all stages $j \in \{1,2,..,n\}$. Hence, we only consider the demand bounds $D(\tau_j^*)$ over τ_j^* consecutive periods as the demand bounds applied in the system.

The arrival demand at Stage *n* is the unbounded external demand that occurs at this stage. The satisfied demand at Stage *n* is the demand that can be satisfied with 100% service by the system, i.e. the part of the demand that lies within the demand bounds. At Stage *n*, the satisfied demand $v(t)$ in period *t* can be formulated as a recursive function that integrates the previous satisfied demands, the unbounded arrival demand $\tilde{d}(t)$ in period *t* and the demand bounds applied at each stage $j \in \{1,2,..,n\}$. In what follows, we explain how to formulate the satisfied demand function $v(t)$.

Let assume that Stage *n* provides an immediate service to the external customer (i.e. $s_n^{out} = 0$) and $\tau_n^* > 0$. For an internal or supply stage $j \in \{1,2,..,n-1\}$ with $\tau_j^* > 0$, if satisfied demand over τ_j^* periods exceed $D(\tau_j^*)$ at Stage *n*, stage *j* will be unable to satisfy demand of its downstream stage with 100% service within the quoted guaranteed-service time. Hence, the demand satisfied from safety stock should not exceed $D(\tau_j^*)$ over $\tau_j^* > 0$ for each stage

$j \in \{1,2,..,n\}$ at Stage *n*. The excess demand can be truncated at Stage *n* in order to remedy this issue. Hence, we express the satisfied demand $v(t)$ in period *t* at Stage *n* by the following recursive function:

$$\begin{aligned} &v(0) = \tilde{d}(0) = 0 \\ &v(t) = \min\{\tilde{d}(t), D(\tau_n^*) - v(t-\tau_n^*, t-1), ..., D(\tau_1^*) - v(t-\tau_1^*, t-1)\} \text{ for } t > 0 \end{aligned} \tag{3.2}$$

where $v(a,b)$ is the satisfied demand over the time interval $(a,b]$ with $v(a,b) = 0$ for $a \geq b$ and $v(a,b) = v(0,b)$ for $a < 0$. We note that if $s_n^{out} > 0$ for the demand stage, the corresponding satisfied demand function can be obtained by setting $v(t) = 0$ for $t \leq s_n^{out}$ and by replacing $\tilde{d}(t)$ in (3.2) by $\tilde{d}(t - s_n^{out})$ for $t > s_n^{out}$.

We denote the unbounded arrival demand during τ consecutive periods at Stage *n* by $\tilde{D}(\tau)$. For ease of presentation, we use the notation $V(\tau)$ to denote the satisfied demand during τ at Stage *n*. By definition, the target CSL α verifies the following:

$$\alpha = P\left(\tilde{D}(\tau_j^*) \leq D(\tau_j^*)\right) \forall j : j \in \{1,2,..,n\} \text{ where } \tau_j^* > 0$$

However, the effectively observed CSL at demand Stage *n* is:

$$\alpha' = P\left(\tilde{D}(\tau_n^*) \leq D(\tau_n^*) \wedge \left\{V(\tau_j^* - 1) + \tilde{D}(1) \leq D(\tau_j^*) \ \forall j : j \in \{1,2,...,n-1\} \text{ where } \tau_j^* > 0\right\}\right) \tag{3.3}$$

Clearly, the effectively observed CSL α' at demand Stage *n* is equal to α if there exists a unique demand bound $D(\tau_n^*)$ in the system over τ_n^* consecutive periods. This happens when safety stocks are solely located at demand Stage *n*. The effectively observed CSL α' may be less that the target CSL α if there exist at least two stages with different strictly positive net replenishment times. In this case, there exist at least two stages that hold safety stocks to cover different net replenishment times and the unbounded arrival demand at Stage *n* should be truncated according to the corresponding demand bounds so that both stages guarantee 100% service within the quoted service times. In the latter case, the effectively observed CSL α' (i.e. the percentage of time that safety stocks cover the demand variation) is lower than the target CSL α.

We note that the satisfied demand function expressed in (3.2) and the effectively observed CSL expressed in (3.3) are also valid for assembly systems where each stage has at most one downstream stage. In this case, there will be a single demand stage which can be indexed by *n*. However, these expressions cannot be directly used for networks containing distribution parts in which internal or supply stages have several downstream stages. First, if there exist several demand stages in the network, target CSL values of different demand stages should be taken into account while specifying the demand bounds for internal and supply stages (stage indexes for demand bounds reappears in this case). Second, when the total unbounded arrival demand exceeds the demand bound of an upstream stage having several downstream stages, one should define how the arrival demands will be truncated at these downstream stages. This is similar to the stock allocation problem encountered in installation stock periodic-review policies where downstream stages place orders at the same time (see, e.g., Jackson, 1988; Graves, 1996; Marklund and Rosling, 2012). The upstream stages face the decision of how to allocate their stock to their downstream stages when the total arrival demand exceeds their demand bound. In this case, an allocation rule should be incorporated in expression (3.2) to determine the satisfied demand function at each demand stage.

Clearly, it is not straightforward to analytically calculate the effectively observed CSL α' under the existence of different demand bounds in the system. To the best of knowledge, Minner (2000) is the only one who attempts to express the effectively observed CSL α' and to present the CSL deviation under the GSM setting. This result is based on a three-stage serial system facing Mixed Erlang demand. He obtains α' by calculating equivalent echelon order-up-to levels, i.e. by summing up the optimal installation order-up-to levels S_j^* and using the service level formulas presented by Van Houtum et al. (1996). Van Donselaar (1989) develops similar expressions for two-stage serial systems. The formulas of Van Donselaar (1989) and Van Houtum et al. (1996) are derived under the SSM assumptions, i.e. in case of unbounded demand and demand backordering. Indeed, under demand truncations and guaranteed-service times these formulas may be inappropriate to express α'. Section

3.1.2 presents an illustrative example and provides an appropriate approximation to calculate the value of α'.

3.1.2 An Illustrative Example and an Approximation

In order to illustrate findings related to Section 3.1.1, we consider an example of two-stage serial supply chain system (see Figure 3.1). The unbounded arrival demand at Stage 2 is normally distributed with an average $\mu = 10$ and a standard deviation $\sigma = 3$ per period. The number of input items required from upstream Stage 1 to obtain one output item at downstream Stage 2 is equal to 1 $(\theta_{12} = 1)$. We assume immediate service towards external customer, i.e. the service time quoted by Stage 2 is equal to 0. The optimal solution obtained by solving *P*0 leads to the net replenishment times $\tau_1^* = 2$ and $\tau_2^* = 1$ at Stage 1 and 2, respectively. In accordance with the GSM literature, we use (3.1) to specify the demand bounds at Stage 1 and 2. The target CSL α used to specify these demand bounds leads to the safety factor $z = \Phi^{-1}(\alpha)$. Hence, demand satisfied from safety stock is assumed to be bounded by $D(2)$ at Stage 1 over two consecutive periods and by $D(1)$ at Stage 2 over one period where $D(2) = 2\mu + z\sigma\sqrt{2}$ and $D(1) = \mu + z\sigma$. The corresponding installation order-up-to levels are: $S_1^* = D(2)$ and $S_2^* = D(1)$.

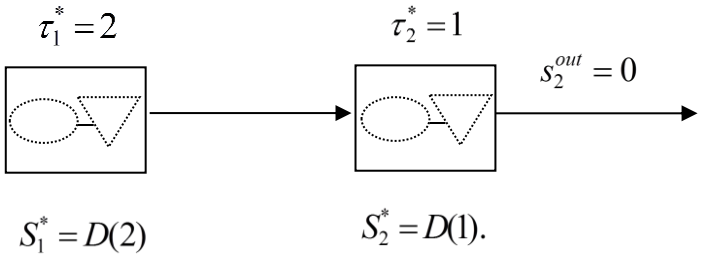

Figure 3.1: The two-stage serial system example

Under the GSM setting, the maximum amount that Stage 2 can replenish from Stage 1 during two consecutive periods is equal to *D*(2). Hence, at Stage 2, the maximum amount of demand which can be satisfied from safety stocks during two consecutive periods should be truncated by *D*(2). Otherwise, Stage 1 cannot guarantee a 100% service level to Stage 2 within its specified service time.

For this example, the satisfied demand at Stage 2 in period $t \geq 0$ can be expressed as follows:

$$v(0) = \tilde{d}(0)$$

$$v(1) = \min\left\{\tilde{d}(1), D(1)\right\} \tag{3.4}$$

$$v(t) = \min\left\{\tilde{d}(t), D(1), D(2) - v(t-1)\right\} \text{ for } t \geq 2$$

The CSL of this system is measured at Stage 2 where the cycle is equal to one period. Hence, the effectively observed CSL α' at Stage 2 is equivalent to the probability that the satisfied demand $v(t)$ is equal to the unbounded arrival demand $\tilde{d}(t)$ in period *t*:

$$\alpha' = P\left(\tilde{D}(1) \leq D(1) \wedge V(1) + \tilde{D}(1) < D(2)\right) = P\left(\tilde{d}(t) = v(t)\right) \tag{3.5}$$

The service level formulas used by Minner (2000) to calculate α' becomes the following for the considered example:

$$\alpha^{Minn} = P\left(\tilde{D}(1) \leq D(1) \wedge \tilde{D}(3) \leq D(1) + D(2)\right) = \int_0^{D(1)} \int_0^{D(1)+D(2)-u} f_2(v) f_1(u) dv du \tag{3.6}$$

where $f_\tau(.)$ is the τ-period demand probability density function. The CSL proposed by Minner (2000) α^{Minn} does not consider the truncations of the unbounded arrival demand. He implicitly assumes that the considered system is equivalent to a SSM setting where unbounded arrival demand is propagated through the system, unsatisfied demand is backordered and stages operate with order-up-to policies with echelon order-up-to levels $\hat{S}_1^* = D(1) + D(2)$ and $\hat{S}_2^* = D(1)$. However, this does not reflect the setting that we consider under the GSM assumptions.

For this example, in order to provide a more approximate value α^{Appr} for the effectively observed CSL α', we propose the following expression:

$$\alpha^{Appr} = F_1(D(2) - D(1)) + \int_{D(2)-D(1)}^{D(1)} \int_0^{D(2)-u} f_1(v) f_1(u) dv du \tag{3.7}$$

where $F_1(.)$ is the cumulative (normal) distribution function of one-period demand. In (3.7), the first term is the probability that the arrival demand does not exceed $D(2) - D(1)$

in an arbitrary period. The second term is the probability that the arrival demand exceeds $D(2)-D(1)$ in an arbitrary period but does not exceeds $D(2)$ during two consecutive periods. This represents an approximation as the second term considers the unbounded arrival demand during two consecutive periods instead of the arrival demand plus the satisfied demand in an arbitrary period. Indeed, the proposed approximation α^{Appr} underestimates the effectively observed CSL α' value.

In order to estimate the effectively observed CSL α', we simulate the considered example generating random values for the unbounded arrival demand in a Microsoft Excel sheet. We calculate the satisfied demand in each period using (3.4). We then assess α' using the probability given in (3.5). We observe that results converge for a simulation length of 100,000 periods.

For different target CSL α values, Table 3.1 summarises α' obtained by simulation and the relative gaps $(\alpha-\alpha')/\alpha$. Simulation results show that the gap between α' and α is really small when α tends to 100%. This gap increases up to a certain point when α decreases. It reaches its maximum value when α is in the interval of (80%, 70%). Then, the gap decreases and reaches zero for $\alpha = 50\%$. Indeed, when $\alpha = 50\%$, z equals zero which makes $D(2) = 2D(1)$. Since the cycle length equals one period for this example, the probability that the satisfied demand is equal to the arrival demand in period *t* becomes 50% under demand bounds $D(1)$ and $D(2)$.

Table 3.2 presents α^{Minn} obtained by (3.6), α^{Appr} obtained by (3.7), the relative gaps $(\alpha'-\alpha^{Minn})/\alpha'$ and $(\alpha'-\alpha^{Appr})/\alpha'$. The relative gaps measured between the effectively observed CSL α' and α^{Minn} points out that the formula proposed by Minner (2000) cannot correctly estimates α' for low values of target CSL. However, α^{Appr} is quiet appropriate to estimate α' for this example. The average relative gap between α' and α^{Appr} is less than 1.00%. This approximation may be considered as a basis in estimating the effectively observed CSL for more complex systems.

In what follows, we present an extensive numerical analysis that investigates the gap between the effectively observed and the target CSL. The next section (Section 3.1.3) presents the test problems used for this numerical analysis.

α	α'	$(\alpha - \alpha')/\alpha$
99.00%	98.40%	0.61%
95.00%	93.11%	1.99%
90.00%	87.15%	3.17%
85.00%	81.76%	3.81%
80.00%	76.56%	4.30%
75.00%	71.63%	4.49%
70.00%	67.06%	4.20%
65.00%	62.63%	3.65%
60.00%	58.38%	2.70%
55.00%	54.09%	1.65%
50.00%	50.00%	0.00%

Table 3.1: CSL deviation results for the two-stage serial system

α	α'	α^{Minn}	α^{Appr}	$(\alpha' - \alpha^{Minn})/\alpha'$	$(\alpha' - \alpha^{Appr})/\alpha'$
99.00%	98.40%	98.92%	98.38%	-0.53%	0.02%
95.00%	93.11%	94.34%	92.75%	-1.32%	0.39%
90.00%	87.15%	88.07%	86.50%	-1.06%	0.75%
85.00%	81.76%	81.42%	80.80%	0.42%	1.17%
80.00%	76.56%	74.59%	75.52%	2.57%	1.36%
75.00%	71.63%	67.67%	70.58%	5.53%	1.47%
70.00%	67.06%	60.77%	65.95%	9.38%	1.66%
65.00%	62.63%	53.96%	61.60%	13.84%	1.64%
60.00%	58.38%	47.32%	57.50%	18.94%	1.51%
55.00%	54.09%	40.90%	53.64%	24.39%	0.83%
50.00%	50.00%	34.75%	50.00%	30.50%	0.00%

Table 3.2: Results obtained by using different formulas that estimate the effectively observed CSL in the two-stage serial system

3.1.3 Test Problems

The test problems presented in this section are used for the numerical analyses conducted along Chapter 3. In the first part of each numerical analysis, we test the impact of each assumption under different combinations of lead time, stage cost and target customer service level alternatives in a five-stage serial system. In the second part, a real-world five-

echelon assembly system is considered and thus, the consequences associated with each assumption are illustrated in a real-world problem setting.

Five-Stage Serial System

Test problems presented here are inspired from Schoenmeyr (2008). We consider a five-stage serial system (see Figure 3.2) and decreasing, uniform and increasing alternatives for stage costs and lead times (see Table 3.3). The terms "increasing" and decreasing" should be understood in terms of going downstream starting from the supply stage (Stage 1). The stage cost c_j represents the cost added at stage *j*. Hence, the cumulative stage cost at the demand stage (Stage 5) represents the cost of the final product. The per-unit holding cost h_j at stage *j* is determined by multiplying the cumulative stage cost at stage *j* (i.e. the cost of the processed item p_j at stage *j*) by a holding cost rate of $\upsilon = 35\%$. The scalar $\theta_{j-1,j}$ is assumed to be 1 for each $j \in \{1,2,..,5\}$. The maximum service time s_5^{out} at Stage 5 is assumed to be zero. External demand in each period is independent and normally distributed with parameters $(\mu = 10, \sigma = 1)$, $(\mu = 10, \sigma = 3)$ and $(\mu = 30, \sigma = 3)$. The 27 test problems considered corresponds to the permutations of these lead time, stage cost and demand process alternatives.

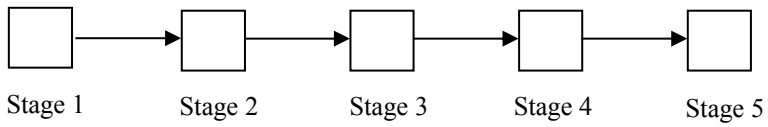

Figure 3.2: Five-stage serial supply chain

Stage Cost / Lead Time	Stages				
	1	*2*	*3*	*4*	*5*
Decreasing	36	28	20	12	4
Uniform	20	20	20	20	20
Increasing	4	12	20	28	36

Table 3.3: Stage cost and lead time alternatives for the five-stage serial system

Real-World Assembly System

This test problem is previously presented by Graves and Willems (2003). It corresponds to a real-world bulldozer assembly and manufacturing supply chain. The relevant network is depicted in Figure 3.3. Table 3.4 provides the stage cost and lead time data for this

system. The scalar $\theta_{i,j}$ equals 1 for each arc $(i, j) \in \mathbf{A}$. The company applies an annual holding cost rate of $\upsilon = 30\%$ when calculating inventory costs (1 year = 260 days). The demand bound at each stage is set using (3.1) based on 95 % target CSL. External demand is assumed to be normally distributed. The average daily demand and the daily standard deviation parameters considered in Graves and Willems (2003) are $\mu = 5$ and $\sigma = 3$, respectively. Clearly, the daily external demand is not well modelled as being from a normal distribution under these parameters. However, the assumption of normality seems plausible for the demand over the net replenishment time of a stage given the range of lead times listed in Table 3.4. We use the assumption of normality for this test problem as it is initially considered by Graves and Willems (2003).

Stage	Stage Cost ($)	Lead Time (days)
Boggie Assembly	575	11
Brake Group	3,850	8
Case	2,200	15
Case & Frame	1,500	16
Chassis/Platform	4,320	7
Common Subassembly	8,000	5
Dressed-out engine	4,100	10
Drive Group	1,550	9
Engine	4,500	7
Fans	650	12
Fender Group	900	9
Final Assembly	8,000	4
Final Drive & Brake	3,680	6
Frame Assembly	605	19
Main Assembly	12,000	8
Pin Assembly	90	35
Plant Carrier	155	9
Platform Group	725	6
Rollover Group	1,150	8
Suspension Group	3,600	7
Track Roller Frame	3,000	10
Transmission	7,450	15

Table 3.4: Stage cost and lead time data for the bulldozer supply chain

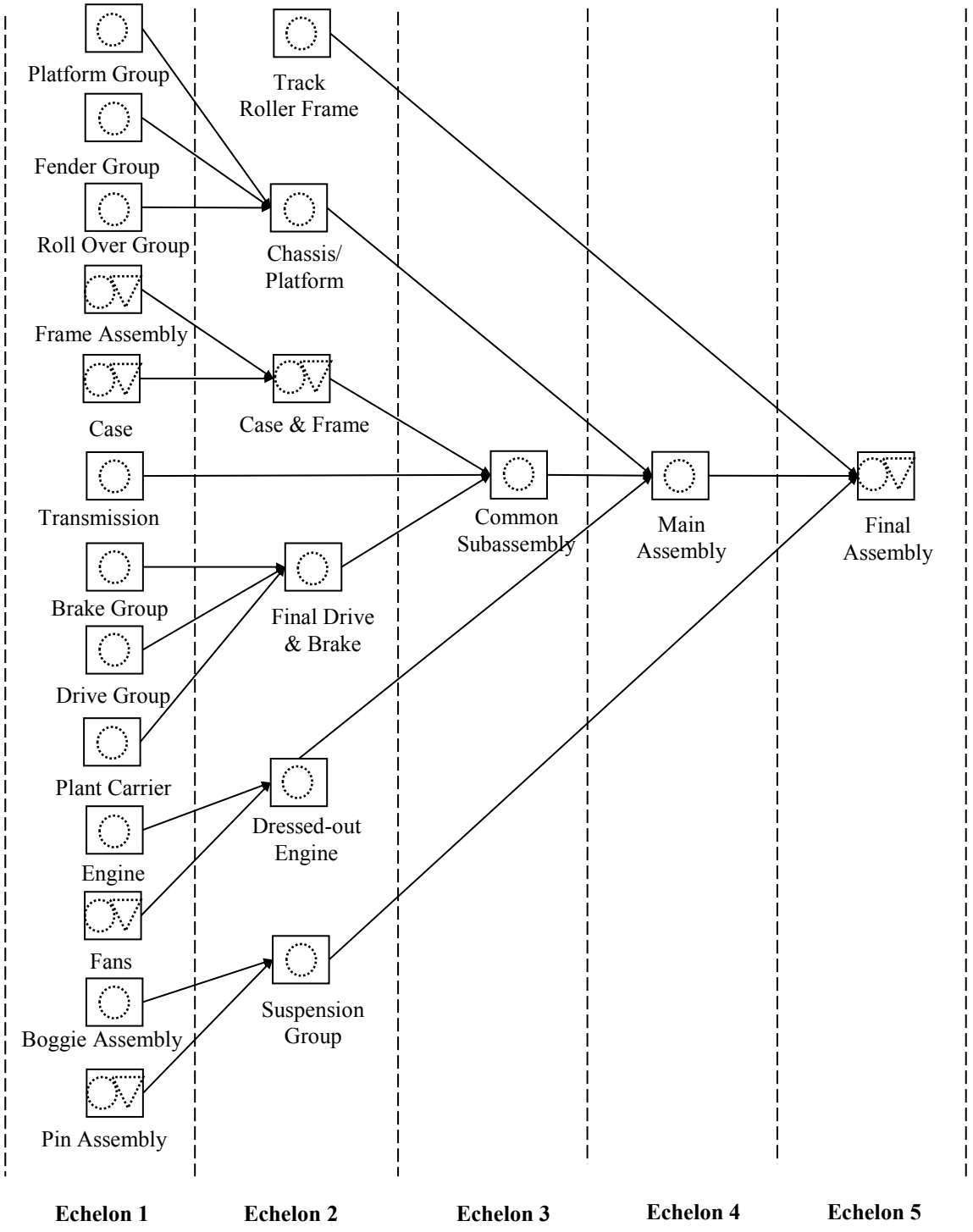

Figure 3.3: The bulldozer supply chain

3.1.4 Numerical Analysis

In this section, simulation experiments are carried out to assess the deviation between the effectively observed and the target CSL for the test problems pertaining to the five-stage serial and the real-world assembly system presented in Section 3.1.3.

Five-Stage Serial System

We determine the optimal solution for each of the 27 test problems presented in Section 3.1.3 by solving the problem *P*0 (given in Section 1.2.2, page 30). We note that the optimal net replenishment times are insensitive to the demand parameters (μ,σ), the safety factor value (z) and the holding cost rate (υ). Given the optimal net replenishment times (see Table 3.5, column 3), we calculate the existing demand bounds using (3.1) based on different safety factor values that relate to target service levels ranging from 50% to 99%. We simulate the five-stage serial system on an Excel sheet by generating random demands that correspond to the considered normal distributions. The simulation length is set to 100,000 periods which is found long enough for convergence. We then determine the satisfied demand in each period using (3.2) and the effectively observed CSL using (3.3).

We observe that the CSL deviations are insensitive to the demand parameters. The relative CSL deviations presented in Table 3.5 are calculated by $(\alpha-\alpha')/\alpha$ for each test problem. In accordance with the findings presented in Section 3.1.1, the CSL at Stage 5 is not affected if the only demand bound applied in the system is the demand bound of Stage 5, i.e. if the safety stock is solely carried at Stage 5. Otherwise, the CSL deviates from its target value.

Differently from the results obtained for the two-stage system example in Section 3.1.2, in most of the cases, the gap between the target and the effectively observed CSL continuously decreases when the target CSL increases. This happens when the net replenishment times of the upstream stage is smaller than that of the demand stage (see the case of increasing stage costs and decreasing lead times for a counterexample). Besides, the CSL at Stage 5 is mostly affected if the net replenishment time of an upstream stage is much smaller than its own net replenishment time (see, e.g., the case of increasing stage costs and increasing lead times). The relative gap between the target and the effectively observed CSL values is 3.4% on average.

Stage Cost	Lead Time	Optimal Net Repl. Times (from Stage 1 to 5)	Target Cycle-Service-Levels				
			80%	*85%*	*90%*	*95%*	*99%*
Decreasing	Decreasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%
	Uniform	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%
	Increasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%
Uniform	Decreasing	(36,0,0,0,64)	3.29%	2.61%	1.91%	1.08%	0.33%
	Uniform	(20,0,0,0,80)	5.13%	4.03%	2.92%	1.65%	0.41%
	Increasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%
Increasing	Decreasing	(36,28,20,0,16)	1.12%	1.08%	1.01%	0.85%	0.35%
	Uniform	(20,20,0,0,60)	4.83%	3.86%	2.80%	1.57%	0.38%
	Increasing	(4,12,0,0,84)	12.79%	10.27%	7.46%	4.25%	1.08%

Table 3.5: CSL deviation results for the five-stage serial problems

Stage Cost	Lead Time	Optimal Net Repl. Times (from Stage 1 to 5)	Target Cycle-Service-Levels					
			50%	*55%*	*60%*	*65%*	*70%*	*75%*
Decreasing	Decreasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Uniform	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	Increasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Uniform	Decreasing	(36,0,0,0,64)	6.55%	6.00%	5.54%	4.98%	4.42%	3.85%
	Uniform	(20,0,0,0,80)	10.81%	9.95%	8.98%	8.13%	7.13%	6.17%
	Increasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Increasing	Decreasing	(36,28,20,0,16)	0.90%	0.94%	0.99%	1.06%	1.08%	1.11%
	Uniform	(20,20,0,0,60)	10.02%	9.15%	8.27%	7.51%	6.54%	5.73%
	Increasing	(4,12,0,0,84)	26.36%	24.09%	21.80%	19.73%	17.46%	15.09%

Table 3.5: CSL deviation results for the five-stage serial problems (continued)

Real-World Assembly System

In this section, we consider the bulldozer assembly and manufacturing supply chain presented in Section 3.1.3. The optimal safety stock placements obtained by solving the problem *P*0 is illustrated graphically in Figure 3.3. A triangle within a stage designates that the stage holds safety stock. The optimal net replenishment times are given for stages that hold safety stock in Table 3.6 (the net replenishment time of other stages equals zero). We note that the optimal net replenishment times are independent of the target CSL applied in the system.

As shown in Table 3.6, demand bounds of internal and supply stages are the demand bounds of 1, 2, 14, 15 and 19 consecutive periods. The demand bound associated with the demand stage corresponds to longest net replenishment time in the system. Using the

parameters given in Section 3.1.3, we generate random values for the normally distributed arrival demand on an Excel sheet over an interval of 100,000 periods (which is sufficiently large for convergence). In each period, if the arrival demand exceeds a demand bound applied in the system, we truncate the excess demand using (3.2). Then, we assess the effectively observed CSL α' using (3.3).

Stage	Optimal Net Repl. Times (days)
Case	15
Case & Frame	1
Fans	2
Final Assembly	32
Frame Assembly	19
Pin Assembly	14

Table 3.6: Optimal net replenishment times

Table 3.7 presents the results obtained for different target CSL (α) values. These results show that the gap between the effectively observed and the target CSL may be significant in a real-world system. Among different target CSL values from 50% to 99%, the relative CSL deviation is 25% on average. In this example, differently from the results obtained for the two-stage system example in Section 3.1.2, the CSL deviation continuously increases when the target CSL decreases. We note that this happens when the net replenishment times associated with the upstream stages are smaller than that of the demand stage.

α	α'	$(\alpha-\alpha')/\alpha$
99.00%	96.88%	2.14%
95.00%	87.67%	7.71%
90.00%	78.48%	12.80%
85.00%	70.02%	17.62%
80.00%	62.45%	21.94%
75.00%	55.54%	25.94%
70.00%	48.92%	30.12%
65.00%	42.89%	34.02%
60.00%	37.46%	37.56%
55.00%	32.45%	41.00%
50.00%	27.79%	44.43%

Table 3.7: CSL deviation results for the bulldozer supply chain

3.2 Guaranteed-Service Times Assumption

In the GSM, each stage quotes a guaranteed-service time to its internal/external customers. Under this assumption, there are no backorders between customer-supplier stages. Each stage guarantees to satisfy the realised demand with 100% service within its guaranteed-service time. The relaxation of this assumption will cause the variability of deliveries, i.e. stages will experience backorders due to occasional stock-outs at their upstream stages. However, by relaxing this assumption for internal and supply stages (non-demand stages), we can find a better solution in terms of cost that still ensures the external customer service requirements. To the best of knowledge, the work of Graves and Willems (2000) contains the only analysis available in the literature that quantifies the impact of the guaranteed-service time assumption. For serial systems, they formulate a model without the guaranteed-service time assumption at non-demand stages and provide a limited computational study to measure the cost performance of this model. In this section, we extend their results to assembly systems and provide a more detailed numerical analysis.

In what follows, we first present the model which relaxes the guaranteed-service time assumption for internal and supply stages (Section 3.2.1). We then examine the cost impact of this assumption by comparing the total safety stock costs obtained by using the original GSM (with the guaranteed-service time assumption at all stages) and the relaxed model (without the guaranteed-service time assumption at non-demand stages). The numerical analysis is carried out for five-stage serial systems and for a real-world assembly system (Section 3.2.2).

3.2.1 Relaxed Model

In the relaxed model, all assumptions of the original model are maintained except the guaranteed service time assumption for non-demand stages. The guaranteed-service times exist only towards external customers (at demand stages). Hence, upstream stages do not quote guaranteed-service times to their internal customers and unsatisfied demands at internal and supply stages are backordered. There are no constraints on internal service levels; rather, internal service levels depend on order-up-to levels which are chosen to minimise the total

safety stock cost of the system. We use this model to compare the performance of the solutions obtained with and without the guaranteed-service time assumption at non-demand stages. In what follows, we first present the model for serial systems then we show how to extend the relevant results to assembly systems.

Solution for Serial Systems

In this section, we consider an *n*-stage serial system for which the optimal policy of the relaxed model is provided by Graves and Willems (2000). Stages are increasingly numbered from upstream to downstream. Without loss of generality, the service time quoted to the external customer is assumed to be zero ($s_n^{out} = 0$) and the coefficient of utilisation $\theta_{j-1,j} = 1$ for each $j \in \{2,...,n\}$. As in the original GSM, each stage operates with a periodic-review, (installation) order-up-to policy, i.e. at the beginning of each period, each stage observes the external demand and places an order for this amount. We denote the external demand in period t by $d(t)$, the external demand over $(a,b]$ by $d(a,b)$ and the demand upper bound during τ periods by $D(\tau)$. We note that the on-hand inventory of stage $j \in \{1,2,...,n\}$ at the end of period t is the positive value of its net inventory where $(I_j(t))^+ = \max\{I_j(t),0\}$. The backorder $B_j(t)$ at the end of period t is the amount that has been ordered by stage j's internal customer but not yet delivered. This is equivalent to the negative part of the net inventory, $B_j(t) = (-I_j(t))^+$. We assume that at time $t = 0$, the on-hand inventory level at stage $j \in \{1,2,...,n\}$ is equal to its order-up-to level S_j, $(I_j(0))^+ = S_j$, and its backorder level is equal to zero, $B_j(0) = 0$. For $j \in \{1,2,...,n\}$, the on-hand inventory $(I_j(t))^+$ and the backorder $B_j(t)$ at the end of period t can be expressed as:

$$(I_j(t))^+ = (S_j - d(t - L_j, t) - B_{j-1}(t - L_j))^+ \tag{3.8}$$

$$B_j(t) = (d(t - L_j, t) + B_{j-1}(t - L_j) - S_j)^+ \tag{3.9}$$

From (3.8) and (3.9), the net inventory level at stage $j \in \{1,2,...,n\}$ is:

$$I_j(t) = (I_j(t))^+ - B_j(t) = S_j - d_j(t - L_j, t) - B_{j+1}(t - L_j)$$

We can write the expected on-hand inventory level at stage j as follows:

$$E[(I_j(t))^+] = S_j - \mu_j L_j + E[B_j(t)] - E[B_{j+1}(t)] \tag{3.10}$$

Hence, the total safety stock cost becomes:

$$\sum_{j=1}^{n} h_j \left\{ S_j - \mu_j L_j \right\} - \sum_{j=1}^{n-1} h_{j+1}^e E[B_j(t)] \tag{3.11}$$

where $h_j^e = h_j - h_{j-1}$ is the echelon holding cost. We assume that the holding costs do not decrease from upstream to downstream, i.e. $h_j^e \geq 0$ for all $j \in \{2,...,n\}$.

As in the original model, we assume that external suppliers provide 100% service which makes $B_0(t) = 0$ by definition. From (3.9), $B_j(t)$ becomes:

$$\begin{aligned} B_j(t) = \max\{ 0, & d(t - L_j, t) - S_j, d(t - L_j - L_{j-1}, t) - S_j - S_{j-1}, ..., \\ & d(t - L_j - L_{j-1} - ... - L_1, t) - S_j - S_{j-1} - ... - S_1 \} \end{aligned} \tag{3.12}$$

In order that Stage *n* provides 100% service level to the external customer for demand within the demand bounds, $B_n(t)$ should be equal to 0 for all *t*. Hence, the order-up-to levels should satisfy the following constraints:

$$S_n + S_{n-1} + ... + S_j \geq D(L_n + L_{n-1} + .. + L_j) \qquad \forall\ j \in \{1,2,..,n\} \tag{3.13}$$

We define the relaxed model *P*1 which minimises (3.11) subject to (3.13) and non-negativity constraints as:

$$P1: \min \sum_{j=1}^{n} h_j \left\{ S_j - \mu_j L_j \right\} - \sum_{j=1}^{n-1} h_{j+1}^e E[B_j(t)]$$

$$\text{s.t. } S_n + S_{n-1} + ... + S_j \geq D(L_n + L_{n-1} + .. + L_j) \qquad \forall\ j \in \{1,2,..,n\}$$

$$S_j \geq 0 \qquad \forall\ j \in \{1,2,..,n\}$$

Graves and Willems (2000) prove that an optimal solution to *P*1 satisfies all the constraints in (3.13) as equalities. Hence, optimal order-up-to levels can be determined by:

$$\begin{aligned} S_n^* &= D(L_n) \\ S_j^* &= D(L_n + ... + L_j) - D(L_n + ... + L_{j+1}) \qquad \forall j \in \{1,2,...,n-1\} \end{aligned} \tag{3.14}$$

It can be shown that the total cost that corresponds to the solution obtained by (3.14) is equal to or less than that of any feasible solution of *P*1. We note that the optimal order-up-to levels do not depend at all on holding costs.

Solution for Assembly Systems

In this section, we will show how the results presented by Graves and Willems (2000) for serial systems can be generalised to assembly systems using the transformation given by Rosling (1989).

We consider an assembly system having *n* stages. We assume that the demand stage is indexed by *n*. We denote the total lead time for stage *j* and all its downstream stages by W_j :

$$\begin{aligned} W_n &= L_n \\ W_i &= L_i + W_j \quad \text{for } i:(i,j) \in \mathbf{A} \end{aligned}$$

The internal and supply stages are indexed so that:

$$\begin{aligned} W_j &\geq W_{j+1} \qquad \forall j \in \mathbf{N_S} \cup \mathbf{N_I} \\ j &> i \qquad\qquad \forall (i,j) \in \mathbf{A} \end{aligned}$$

According to results presented in Rosling (1989), the optimal policies of the assembly system are equivalent to those of a serial system where stage *j*'s immediate downstream stage is stage $j+1$ and the lead time L_j^{ser} of stage *j* in the equivalent serial system is:

$$\begin{aligned} L_n^{ser} &= L_n \\ L_j^{ser} &= W_j - W_{j+1} \qquad \forall j \in \{1,2,...,n-1\} \end{aligned}$$

Using (3.14), the optimal (installation) order-up-to levels for the equivalent serial system (S_j^{ser}) become:

$$\begin{aligned} S_n^{ser} &= D(L_n^{ser}) \\ S_j^{ser} &= D(L_n^{ser} + ... + L_j^{ser}) - D(L_n^{ser} + ... + L_{j+1}^{ser}) \qquad \forall j \in \{1,2,...,n-1\} \end{aligned} \tag{3.15}$$

The optimal (installation) order-up-to levels for the considered assembly system can be determined by:

$$
\begin{aligned}
&S_n^* = S_n^{ser} \\
&S_j^* = S_j^{ser} + S_{j+1}^{ser} + ... + S_{j+k}^{ser} \qquad \forall j \in \{1,2,,..,n-1\} \text{ and } W_j - W_{j+k} = L_j
\end{aligned}
\tag{3.16}
$$

Equations (3.15) and (3.16) lead to the following:

$$
\begin{aligned}
&S_n^* = D(L_n) \\
&S_i^* = D(L_n + ... + L_i) - D(L_n + ... + L_j) \qquad \forall i : (i, j) \in \mathbf{A}
\end{aligned}
\tag{3.17}
$$

Therefore, we can determine the optimal order-up-to levels for an assembly system using (3.17). These results can be extended to distribution systems by considering the stock allocation issue encountered in periodic-review policies.

3.2.2 Numerical Analysis

In this numerical analysis, we compare the cost performance of solutions obtained using the original GSM *P*0 (given in Section 1.2.2, page 30) and the relaxed model *P*1. We use the test problems associated with the five-stage serial and the real-world assembly system presented in Section 3.1.3.

Five-Stage Serial System

We consider the five-stage serial system presented in Figure 3.2 under the lead time and stage cost alternatives given in Table 3.3 (Section 3.1.3), normally distributed demand with parameters $(\mu = 10, \sigma = 1)$, $(\mu = 10, \sigma = 3)$ and $(\mu = 30, \sigma = 3)$ and 90%, 95%, 99% target CSL values. By evaluating all combinations (including the target CSL alternatives) we have a total of 81 test problems. We note that in (3.11), $E[B_j(t)]$ is a nonlinear function of order-up-to levels $S_1, S_2, .., S_j$. Exact values for normally distributed demand are not easy to compute for this function. In this study, we use simulation in order to evaluate $E[B_j(t)]$ for the considered test problems. We set the simulation length sufficiently large for convergence (100,000 periods).

Our numerical analysis enhances results provided by Graves and Willems (2000). They consider a set of 36 test problems pertaining to a three-stage serial system facing Poisson demand under different alternatives of lead time, holding cost, demand rate and target CSL. They show that the safety stock cost for the model with guaranteed-service times

is on average 26% higher than that for the model without this assumption; the range is between 7% and 43%. They show that the size of the gap is insensitive to the choice of demand process and the target CSL. However, the gap becomes larger as the lead time at Stage 5 increases and as the stage cost at Stage 5 increases. They show that the impact on the total inventory cost is less dramatic. The total inventory cost corresponds to the sum of the safety stock cost and the pipeline stock cost. The difference in total inventory costs is 4% on average, with a range from less than 1% to 14%. The gap increases as the pipeline stock cost decreases, namely as the lead time at Stage 5 decreases and as the demand rate decreases.

For the five-stage serial test problems presented in Section 3.1.3, Table 3.8 summarises the relative safety stock cost gaps $(Obj_0 - Obj_1)/Obj_1$ where Obj_0 is the total safety stock cost obtained by solving $P0$ and Obj_1 is the one obtained by solving $P1$. These results show that the total safety stock cost under the guaranteed-service assumption is 35.9% higher on average than the one without this assumption. Hence, the safety stock cost increase becomes more significant for the considered test problems compared to the results obtained by Graves and Willems (2000).

Stage cost	**Lead time**	**Target Cycle-Service-Levels**		
		90%	*95%*	*99%*
Decreasing	Decreasing	34.23%	34.60%	34.06%
	Uniform	21.51%	20.87%	19.48%
	Increasing	11.84%	10.70%	9.26%
Uniform	Decreasing	57.83%	58.23%	57.16%
	Uniform	39.47%	38.42%	36.05%
	Increasing	24.88%	23.04%	20.55%
Increasing	Decreasing	59.07%	59.47%	57.63%
	Uniform	46.97%	45.39%	41.80%
	Increasing	38.52%	35.70%	31.76%

Table 3.8: Relative gaps between the total safety stock costs obtained by solving $P0$ and $P1$

Similarly to Graves and Willems (2000), our results confirm that the safety stock cost gap increases if the stage cost at Stage 5 increases. However, contrary to the results of Graves and Willems (2000), we observe that the gap increases if the lead time at Stage 5 decreases. Besides, we cannot say the gap is insensitive to the target CSL levels for the considered test

problems; the gap usually decreases when the target CSL increases. Additionally, our results show that the safety stock gap is insensitive to demand variability.

We calculate the expected pipeline inventory cost at stage j with $PS_j = h_j L_j \mu$. Similarly to Graves and Willems (2000), we notice that the impact of guaranteed-service time assumption is less significant when the total inventory cost is considered. We observe that the total cost gap depends on the coefficient of variation (σ / μ) and on the target CSL value. Table 3.9 summarises the relative inventory cost gaps for different coefficient of variations averaging the solutions obtained for different target CSL values. This shows that the total cost gap increases when the coefficient of variation increases and lead time at Stage 5 decreases.

Stage Cost	**Lead Time**	**Coefficient of Variation**	
		0.1	*0.3*
Decreasing	Decreasing	0.69%	1.99%
	Uniform	0.38%	1.10%
	Increasing	0.18%	0.52%
Uniform	Decreasing	1.31%	3.74%
	Uniform	0.77%	2.21%
	Increasing	0.41%	1.19%
Increasing	Decreasing	1.69%	4.79%
	Uniform	1.03%	2.94%
	Increasing	0.68%	1.97%

Table 3.9: Average relative gaps between the total inventory costs obtained by solving $P0$ and $P1$

Real-World Assembly System

In this section, we consider the real-world system presented in Section 3.3.2. We obtain the optimal order-up-to level of each stage for the relaxed model $P1$ using (3.17). The backorder levels at internal stages are evaluated by simulation considering these order-up-to levels. Resulting safety stock costs are displayed in Table 3.10. Results obtained for this system shows that the cost increase due to the guaranteed-service time assumption may be significant for a real-world system. This example represents 42.2% safety stock cost increase for which the total stock cost increase including the pipeline stock cost is 7.0%.

Stage	Safety Stock Cost Obtained Solving $P0$ ($)	Safety Stock Cost Obtained Solving $P1$ ($)	Pipeline Stock Cost ($)
Boggie Assembly	-	1,308	9,488
Brake Group	-	6,018	46,200
Case	12,614	4,823	49,500
Case & Frame	6,373	10,906	103,320
Chassis/Platform	-	10,678	74,498
Common Subassembly	-	31,214	217,425
Dressed-out engine	-	18,677	138,750
Drive Group	-	2,627	20,925
Engine	-	6,782	47,250
Fans	1,361	1,407	11,700
Fender Group	-	1,695	12,150
Final Assembly	607,969	190,034	435,600
Final Drive & Brake	-	10,670	83,115
Frame Assembly	3,904	1,579	17,243
Main Assembly	-	103,958	688,020
Pin Assembly	499	483	4,725
Plant Carrier	-	263	2,093
Platform Group	-	1,053	6,525
Rollover Group	-	2,004	13,800
Suspension Group	-	8,701	44,783
Track Roller Frame	-	8,743	45,000
Transmission	-	21,238	167,625
TOTAL	632,719	444,860	2,239,733

Table 3.10: Results obtained by solving $P0$ and $P1$ for the bulldozer supply chain

3.3 Common Review Periods Assumption

In this section, we investigate the impact of the common review periods assumption on the total safety stock cost. In the original GSM, each stage operates with a periodic-review, order-up-to policy with a common review period of one period length. In practice, it may be more appropriate to review the stock status (and to place orders) less frequently than each period considering the economies of scale and/or the availability of resources. Besides, review periods (reorder intervals) to be applied at different stages may be different from each other. This requires introducing stage-dependent review periods. A simple approach to integrate stage-dependent review periods into the GSM would be to aggregate the review period of each stage into its lead time. Hence, one can use the original model $P0$ to obtain an

approximate solution. In this section, we evaluate the safety stock cost increase under such an approximation by comparing it with a model that we propose, that incorporate stage-dependent review periods into the GSM.

As presented in Chapter 2, Bossert and Willems (2007) provide the only work that extends the GSM to allow stage-dependent review periods. They assume that the review period of each stage has an arbitrary and integer value. Under this assumption, although the external demand processes are stationary, the internal demands might turn into cyclic processes. Besides, while expressing the cyclic inventory dynamics, not only the cycle length but also the staggering of internal orders should be taken into account. For instance, two stages that order every week might order on different days of the week. This complicates the evaluation of expected inventory levels at internal stages. Bossert and Willems (2007) provide an approximation for this case.

In this section, differently from the work of Bossert and Willems (2007), we particularly focus on nested review periods for which we provide accurate formulations for expected inventory and safety stock levels. Under a nested policy, every replenishment epoch of an upstream stage coincides with a shipment epoch towards its downstream stage. Indeed, stationary nested policies are proven to be optimal for deterministic serial (see Schwarz, 1973) and assembly systems (Muckstadt and Roundy, 1993). However, they may be suboptimal for distribution and general multi-echelon structures (Roundy, 1985b). Nevertheless, many researchers assume nested policies for distribution and general multi-echelon systems because of their significant practical and computational advantages (see, e.g., Maxwell and Muckstadt, 1985; Yao and Wang, 2006).

We first show how to express the expected inventory and safety stock levels under nested stage-dependent review periods. Incorporating the proposed formulation into the GSM, we develop a model that relaxes the common review periods assumption (Section 3.3.1). We then compare solutions obtained by the original and the relaxed model using the five-stage serial and the real-world assembly test problems presented in Section 3.1.3 under different review period profiles (Section 3.3.2).

3.3.1 Relaxed Model

The original GSM does not model the case where review periods are stage-dependent. In this section, we first present the demand propagation and inventory dynamics in serial systems under nested stage-dependent review periods. Then, we give the corresponding mathematical programming formulation. At the end of this section, we show how to extend this model to general acyclic systems.

Inventory Dynamics for Serial Systems

We consider an *n*-stage serial system where stages are increasingly numbered from upstream to downstream. Without loss of generality, we assume that the coefficient of utilisation $\theta_{j-1,j} = 1$ for each $j \in \{2,...,n\}$. The external demand $d(t)$ is propagated through the system where the realised demand over $(a,b]$ at different stages is represented by $d(a,b)$. The expression presented by Bossert and Willems (2007) leads to the following net inventory balance equation for $R_j \geq 1$ at stage $j \in \{1,2...,n\}$:

$$I_j(t) = S_j - d(t - s_j^{in} - L_j - x_j(t), t - s_j^{out}) \tag{3.18}$$

where $x_j(t)$ reflects the inventory exposure because of review periods. The value of $x_j(t)$ cycles in $[0,R_j)$. The first argument of $d(.,.)$, $t - s_j^{in} - L_j - x_j(t)$ corresponds to the last demand replenished by the stage's upstream stage by period *t*. The second argument of $d(.,.)$, $t - s_j^{out}$ corresponds to the last demand fulfilled by period *t*. In what follows, we express the expected net inventory and safety stock levels for demand and non-demand stages.

Let consider the demand stage (Stage *n*) where the external demand is a continuous-time process. For this stage, the net inventory level $I_n(t)$ is of the saw-tooth form (see Figure 3.4). Without loss of generality, we assume that Stage *n* places orders at times mR_n where *m* is a non-negative integer. In this case, $x_n(t)$ becomes:

$$x_n(t) = (t - s_n^{in} - L_n) \bmod R_n$$

The inventory exposure is the largest when $x_n(t) = R_n - \varepsilon$ where ε is a very small positive number. If $t - s_n^{in} - L_n - R_n < t - s_n^{out}$, Stage *n* should have the amount of inventory to

cover the demand over an interval of length $\tau_n = s_n^{in} + L_n + R_n - s_n^{out}$. In order that Stage *n* provides 100% service to its customers, the base stock level S_n should be set to the demand upper bound $D(\tau_n)$ during the net replenishment time τ_n:

$$S_n = D(s_n^{in} + L_n + R_n - s_n^{out}) \tag{3.19}$$

In (3.19) one can specify the demand bound function as in (3.1). Under the bounded demand assumption, (3.19) makes the net inventory level $I_n(t)$ always positive. Hence, the net inventory equals the on-hand inventory.

Since $x_n(t)$ cycles in $[0, R_n)$ with a cycle length of R_n, the expected demand during this interval can be expressed as:

$$E\left[d(t - s_n^{in} - L_n - x_n(t), t - s_n^{out})\right] = \mu\left(s_n^{in} + L_n + R_n - s_n^{out}\right) - \frac{R_n \mu}{2} \tag{3.20}$$

Using (3.19) and (3.20), we can represent the expected inventory level $E\left[I_n(t)\right]$ as:

$$E\left[I_n(t)\right] = D\left(s_n^{in} + L_n + R_n - s_n^{out}\right) - \mu\left(s_n^{in} + L_n + R_n - s_n^{out}\right) + \frac{R_n \mu}{2} \tag{3.21}$$

We note that (3.21) is similar to the approximation provided by Hadley and Whitin (1963) where there are two components of stock, the safety stock SS_n and the cycle stock CS_n:

$$SS_n = D\left(s_n^{in} + L_n + R_n - s_n^{out}\right) - \mu\left(s_n^{in} + L_n + R_n - s_n^{out}\right)$$

$$CS_n = \frac{R_n \mu}{2}$$

Let now consider a non-demand stage $j \in \{1,2,...,n-1\}$. Under nested review periods R_j / R_{j+1} is an integer for $j \in \{1,2,...,n-1\}$ and it is always possible to synchronise replenishments and expeditions. We assume that the reorder epochs are offset by ω_j to allow each stage $j \in \{1,2,...,n-1\}$ to replenish from its immediate upstream stage at the exact moment an order arrives (becomes available) at the upstream stage. We assume that stage *j* places orders at times $\omega_j + mR_j$ where $\omega_j \in \{1,2,..., R_j - 1\}$. Hence, $x_j(t)$ becomes:

$$x_j(t) = (t - s_j^{in} - L_j - \omega_j) \bmod R_j \qquad \text{for } j \in \{1,2,...,n-1\}$$

Since orders are placed periodically, demand occurs in discrete-time periods at non-demand stages. We note that the inventory level $I_j(t)$ at stage $j \in \{1,2,...,n-1\}$ is not of the saw-tooth form (see Figure 3.5). The inventory exposure is the largest when $x_j(t) = R_j - 1$ and smallest when $x_j(t) = 0$. If $t - s_j^{in} - L_j - R_j - 1 < t - s_j^{out}$, stage j should have the amount of inventory to cover the demand over an interval of length $\tau_j = s_j^{in} + L_j + R_j - 1 - s_j^{out}$.

Since the replenishment of stage j is synchronised by the expedition towards stage $j+1$, the number of orders η_j placed by stage $j+1$ and observed by stage j during τ_j can be calculated by the following floor function:

$$\eta_j = \left\lfloor \frac{s_j^{in} + L_j + R_j - 1 - s_j^{out}}{R_{j+1}} \right\rfloor$$

Therefore, at stage $j \in \{1,2,...,n-1\}$ the length of the time interval that should be covered by stocks becomes $\eta_j R_{j+1}$.

The order-up-to level at stage $j \in \{1,2,...,n-1\}$ should be set to the demand upper bound during $\eta_j R_{j+1}$:

$$S_j = D(\eta_j R_{j+1}) \qquad \text{for } j \in \{1,2,...,n-1\} \tag{3.22}$$

In (3.22) one can specify the demand bound function as in (3.1). In this setting, the expected inventory level $E[I_j(t)]$ becomes:

$$E[I_j(t)] = D(\eta_j R_{j+1}) - \eta_j R_{j+1}\mu + \left(\frac{R_j}{R_{j+1}} - 1\right)\frac{R_{j+1}\mu}{2} \qquad \text{for } j \in \{1,2,...,n-1\} \tag{3.23}$$

In (3.23), we distinguish two components of stock, the safety stock SS_j and the cycle stock CS_j:

$$SS_j = D(\eta_j R_{j+1}) - \eta_j R_{j+1}\mu \qquad \text{for } j \in \{1,2,...,n-1\}$$

$$CS_j = \left(\frac{R_j}{R_{j+1}} - 1\right)\frac{R_{j+1}\mu}{2} \qquad \text{for } j \in \{1,2,...,n-1\}$$

In the optimisation problem, the cycle stock cost may be ignored since it depends only on input parameters and does not affect the optimisation. We note that this does not mean that the cycle stock cost is not a significant part of the inventory cost.

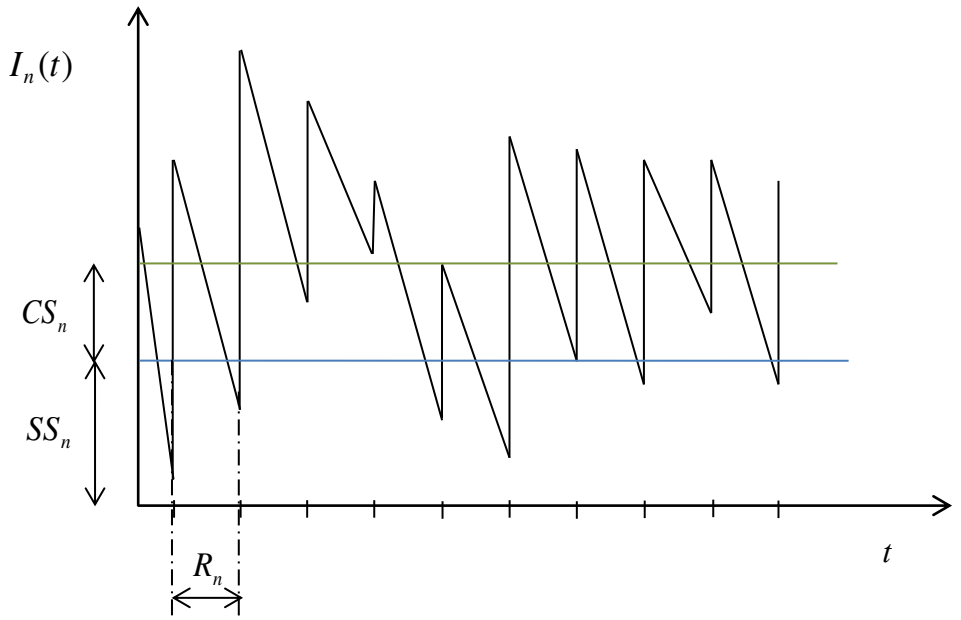

Figure 3.4: Net (on-hand) inventory evolutions for a demand stage

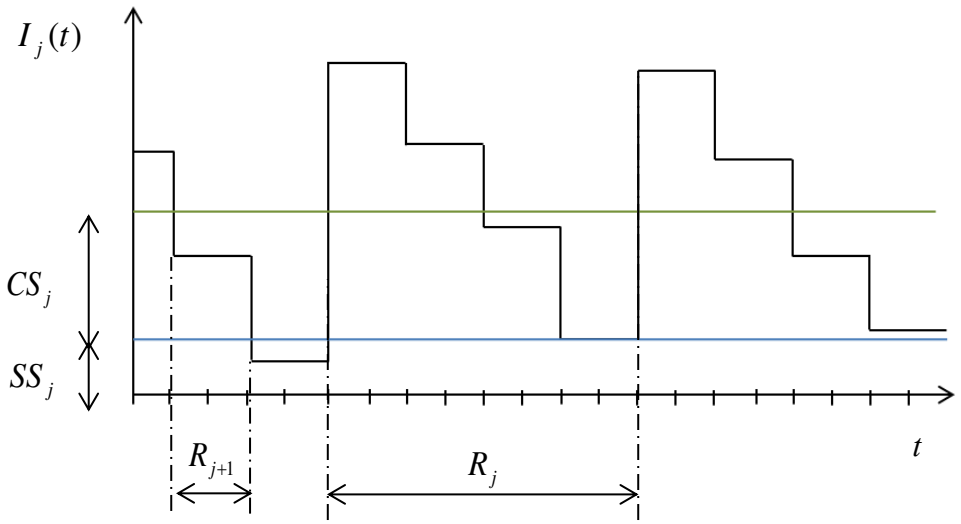

Figure 3.5: Net (on-hand) inventory evolutions for an internal or supply stage

Mathematical Programming Formulation for Serial Systems

The problem of minimising the total safety stock cost in an *n*-stage serial system under nested stage-dependent review periods is formulated as *P*2:

$$P2: \min \sum_{j=1}^{n-1} h_j \left\{ D(\eta_j R_{j+1}) - \eta_j R_{j+1} \mu \right\} + h_n \left\{ D(\tau_n) - \tau_n \mu \right\} \tag{3.24}$$

s.t. $\tau_n = s_n^{in} + L_n + R_n - s_n^{out}$ (3.25)

$\tau_j = s_j^{in} + L_j + R_j - 1 - s_j^{out} \quad \forall j \in \{1,2,...,n-1\}$ (3.26)

$\tau_j - \eta_j R_{j+1} \geq 0 \quad \forall j \in \{1,2,...,n-1\}$ (3.27)

$\tau_j - \eta_j R_{j+1} < R_{j+1} \quad \forall j \in \{1,2,...,n-1\}$ (3.28)

$s_j^{out} \leq s_{j+1}^{in} \quad \forall j \in \{1,2,...,n-1\}$ (3.29)

$s_n^{out} \leq s_n^{client}$ (3.30)

$\tau_j, s_j^{out}, s_j^{in}, \eta_j \geq 0$ and integer $\quad \forall j \in \{1,2,...,n\}$ (3.31)

The difference between *P*0 and *P*2 comes from the definition of the coverage and net replenishment times of stages. Constraints (3.25) and (3.26) define the net replenishment times of stages. Constraints (3.27) and (3.28) give an equivalent representation of the floor function to compute the coverage times of non-demand stages (see also Chapter 4).

Extension to General Acyclic Systems

For assembly systems, the extension of the problem *P*2 is straightforward. Since each stage has at most one downstream stage in an assembly system, index $j+1$ in *P*2 can be replaced by index $k : (j,k) \in \mathbf{A}$. However, if the supply chain network contains a distribution structure, demand realisations from different downstream stages should be differentiated while setting the internal demand bounds. Indeed, in case of several demand stages, the net replenishment time of each demand stage $j \in \mathbf{N_D}$ becomes:

$\tau_j = s_j^{in} + L_j + R_j - s_j^{out} \quad$ for $j \in \mathbf{N_D}$ (3.32)

The demand bound function $D_j(.)$ can be specified using (2.1) (Chapter 2, page 41) for demand stages. Hence, the order-up-to level S_j at stage $j \in \mathbf{N_D}$ can be set to:

$S_j = D_j(s_j^{in} + L_j + R_j - s_j^{out}) \quad$ for $j \in \mathbf{N_D}$ (3.33)

If an internal or supply stage $j \in \mathbf{N_I} \cup \mathbf{N_S}$ has several downstream stages $k:(j,k) \in \mathbf{A}$, we can define the coverage times associated with each downstream stage k as follows:

$$\left\lfloor \frac{s_j^{in} + L_j + R_j - 1 - s_j^{out}}{R_k} \right\rfloor R_k = \eta_{jk} R_k \quad \text{for } j:(j,k) \in \mathbf{A} \tag{3.34}$$

where η_{jk} represents the number of orders placed by stage k observed by stage $j:(j,k) \in \mathbf{A}$ during $\tau_j = s_j^{in} + L_j + R_j - 1 - s_j^{out}$. By adapting the expression given in Graves and Willems (2000) for the demand bounds of non-demand stages one can set the order-up-to level at stage $j \in \mathbf{N_I} \cup \mathbf{N_S}$ as:

$$S_j = \sum_{k:(j,k)\in\mathbf{A}} \theta_{jk} \eta_{jk} R_k \mu_k + \sqrt[p]{\sum_{k:(j,k)\in\mathbf{A}} \left\{ \theta_{jk} \left(D_k(\eta_{jk} R_k) - \eta_{jk} R_k \mu_k \right) \right\}^p} \text{ for } j \in \mathbf{N_I} \cup \mathbf{N_S} \tag{3.35}$$

where $p \geq 1$ is a given constant that relates to the risk pooling effect. Larger values of p correspond to more risk pooling, i.e. larger reduction in demand variability due to combining the demands of multiple downstream stages. Setting $p = 2$ equates to combining standard deviations of independent demand streams. Setting $p = 1$ models the case of no risk pooling.

Therefore, one can extend the problem *P*2 to general acyclic systems by replacing the order-up-to levels in the objective function by (3.33) and (3.35) and by defining the coverage times of non-demand stages by (3.34). We present this extension in more detail in Chapter 4.

3.3.2 Numerical Analysis

This numerical analysis is carried out using the five-stage serial and five-echelon assembly test problems presented in Section 3.1.3. We add another dimension into these test problems by considering different nested review period profiles (see Table 3.11). In particular, we consider uniform, ending and decreasing profiles. Since nested review periods increase from downstream to upstream, these profiles represent feasible alternatives. The lead time of stages given in Table 3.3 and Table 3.4 are considered as the time length $L_j + R_j - 1$ for non-demand stages $j \in \mathbf{N_I} \cup \mathbf{N_S}$ and $L_n + R_n$ for demand stage n. Thus, in *P*0, we consider that the review periods are included into the lead times. We measure the

performance of such an approximation that could be used in the original GSM. In other words, we evaluate the safety stock cost increase due to the lack of modelling of stage-dependent review periods in the original GSM. Optimal solutions are obtained using solver BARON (version 9.3.1) by modelling *P*0 and *P*2 in GAMS 23.7.

Five-Stage Serial System

First, we present the results obtained for the five-stage serial test problems. We denote the safety stock cost obtained by solving the problem *P*2 by Obj_2. Table 3.11 summarises the average relative gaps $(Obj_0 - Obj_2)/Obj_2$ under different review period profiles. We note that the gaps between the solutions obtained by solving *P*0 and *P*2 are insensitive to the demand parameters and the target CSL value (since these parameters turn into a factor of the total safety stock cost).

For the considered test problems, we observe that the total safety stock cost decreases when review periods increase. The gap becomes larger when the review periods are common for all stages (uniform profile). However, starting profiles have no impact on the total safety stock cost. This is due to definition of coverage times at supply stages.

We also notice that the safety stock locations obtained by solving *P*0 and *P*2 may be different for the same input data. When we take into account the demand propagation under stage-dependent review periods, the solution that minimises the total safety stock cost may result in different safety stock locations than the solution obtained by *P*0. For instance, the solution obtained by *P*0 for the test problem that represents the combination of increasing stage costs, increasing lead times and uniform review periods profile (4, 4, 4, 4, 4) corresponds to a solution where safety stocks are held at Stages 1,2 and 5 (see Table 3.5). However, for the same test problem the solution obtained by *P*2 consolidates all safety stocks at Stage 5.

Another interesting result is that the use common review periods across the supply chain may seem reasonable in the sense that the resulting safety stock cost decrease is relatively high for uniform profiles in comparison to starting or decreasing profiles. However, we cannot say which profile is better considering these results. The impact of different review

period profiles on the total safety stock cost must be evaluated together with the associated ordering and cycle stock costs.

Review Periods (from Stage 1 to 5)		Average Safety Stock Cost Gap
Uniform	(2, 2, 2, 2, 2)	2.00%
	(3, 3, 3, 3, 3)	4.32%
	(4, 4, 4, 4, 4)	6.31%
Starting	(2, 1, 1, 1, 1)	0.00%
	(3, 1, 1, 1, 1)	0.00%
	(4, 1, 1, 1, 1)	0.00%
Decreasing	(2, 2, 2, 1, 1)	0.65%
	(3, 3, 3, 1, 1)	1.45%
	(4, 4, 2, 2, 1)	1.92%

Table 3.11: Average relative gaps between the total safety stock costs obtained by solving $P0$ and $P2$ for the five-stage serial system

Real-World Assembly System

Second, we consider the five-echelon real-world assembly system presented in Section 3.1.3. We apply the review period profiles presented in Table 3.11 at different echelons.

Table 3.12 summarises the relative gaps $(Obj_0 - Obj_2)/Obj_2$ between the solutions obtained by solving $P0$ and $P2$. These results show that the safety stock cost increase of using the original model $P0$ under stage-dependent review periods may be significant in a real-world supply chain. Similarly to the previous analysis for the five-stage serial systems, the gap increases when review periods increase. We notice that the gap becomes more important for uniform profiles. This is due to the existence of the floor function term while defining the coverage times of non-demand stages. For instance, the sum of review periods in profile (3, 3, 3, 1, 1) is more than that of (2, 2, 2, 2, 2). However, the cost increase of the later is larger than the former one.

Review Periods (from Echelon 1 to 5)		Safety Stock Cost Gap
Uniform	(2, 2, 2, 2, 2)	5.40%
	(3, 3, 3, 3, 3)	10.93%
	(4, 4, 4, 4, 4)	19.11%
Starting	(2, 1, 1, 1, 1)	0.00%
	(3, 1, 1, 1, 1)	0.00%
	(4, 1, 1, 1, 1)	0.00%
Decreasing	(2, 2, 2, 1, 1)	1.44%
	(3, 3, 3, 1, 1)	3.14%
	(4, 4, 2, 2, 1)	4.46%

Table 3.12: Relative gaps between the total safety stock costs obtained by solving $P0$ and $P2$ for the bulldozer supply chain

3.4 Conclusion

In this chapter, we provided detailed insights on the impacts of the bounded demand, guaranteed-service times and common review periods assumptions.

First, we showed that under the existence of demand bounds, the effectively observed CSL at the most downstream stage of the supply chain may be less than the target CSL. The gap is due to the fact that the effectively observed CSL at the most downstream stage is affected by the demand bounds applied at the upstream stages. The gap between the target and the effectively observed CSL increases when the net replenishment time of an upstream stage is much smaller than that of the most downstream stage. We assessed this deviation by simulation. Simulation studies that we carried out showed that the gap is 25% on average for a real-world assembly system.

Assessing the CSL deviation under the existence of demand bounds over different net replenishment times in the system is not straightforward. We propose an approximation for a simple case with two different demand bounds. Further research can be conducted to propose analytical methods and approximations to calculate this deviation for more general cases. Another important research question that arises from this analysis is how to mitigate the CSL deviation in the GSM setting. Indeed, in order to achieve the target CSL without extraordinary measures at the most downstream stage, the safety factor to be applied at different stages should be appropriately defined. An extension of the GSM can be developed

by considering the safety factors as decision variables of the problem which are subject to a target customer service level constraint. However, finding an optimal solution to this extension may be challenging. In Chapter 5, we propose two different approaches that mitigate the CSL deviation for supply chains operating under the GSM assumptions. Besides, we present some related issues encountered in decentralised supply chains.

Second, we examined the impact of the guaranteed-service time assumption on the total safety stock cost. This analysis enhanced the preliminary results provided by Graves and Willems (2000). We showed that the total safety stock cost under the guaranteed-service time assumption may be significantly higher than the one without this assumption. The total safety stock cost increase may be up to 42.2% for a real-world system. However, in practice, managers seem more comfortable with the notion of guaranteed-service time due to its practical advantages. The guaranteed-service time assumption prevents the variability of deliveries. Indeed, without this assumption, even the inventory control policy for overall system is optimised; a local manager at a certain stage may feel the need of more safety stocks and may deviate from the optimal solution to reduce the variability (Minner, 2000). In order to combine the cost advantage of a solution with backorders and the practical advantage of using guaranteed-service times, one can propose a mixed model where some stages are allowed to backorder demand and some are subject to the guaranteed-service times. We leave this extension as a future work.

Third, we examined the assumption of common review periods. We proposed a model that enables to consider stage-dependent nested review periods. We compared the solution obtained by the proposed model with a simple approximation of the original model that aggregates the review period of stages into their lead times. The numerical analysis showed that this approximation may represent a significant cost increase for long review periods. Besides, the cost increase is relatively large when the review periods are common for all stages. However, we cannot say which review period profile is more appropriate for guaranteed-service supply chains based on these results. The impact of review periods on the total safety stock cost must be evaluated together with the associated fixed ordering costs. Hence, one potential extension of this work is to enable the GSM to simultaneously

determine the optimal review periods (reorder intervals) and safety stock levels (order-up-to levels) by incorporating the fixed ordering cost into the total cost function. We provide this extension in Chapter 4.

In this chapter, the analysis regarding the impacts of the bounded demand, guaranteed service times and common review periods assumptions are conducted individually. Indeed, quantifying the combined effect of these assumptions can be translated as an analysis that compares the GSM and a fairly comparable SSM that excludes the GSM assumptions. We note that such analyses are conducted by Klosterhalfen and Minner (2007; 2010) for serial and two-echelon distribution systems by incorporating the cost of extraordinary measures into the GSM (see also Section 2.3 of Chapter 2).

In this chapter, the results presented for the bounded demand and guaranteed-service times assumptions are limited to serial and assembly systems. The extension of these results to distribution systems requires the consideration of the stock allocation problem encountered in periodic-review policies. We leave the analysis for distribution and general acyclic structures as a future research direction. For the common review periods assumption, we showed how to extend the relaxed model to general acyclic multi-echelon systems. We will recall this result in Chapter 4 while modelling the problem of simultaneously optimising the reorder intervals and order-up-to levels in general acyclic systems.

Appendix: A Remark on Expected Inventory Levels

In this appendix, we provide a remark on the derivation of expected inventory levels under the setting presented in Section 3.1.1.

In the original GSM, the inventory balance equation is expressed as (1.15) (see Chapter 1, page 29). However, under the setting presented in Section 3.1.1, it is accurate to replace the realised demand by the satisfied demand $v_j(t)$ given in (3.2). This leads to the following inventory balance equation:

$$I_j(t) = S_j - v_j(t - s_j^{in} - L_j, t - s_j^{out})$$

Therefore, the expected inventory level $E[I_j(t)]$ becomes:

$$E[I_j(t)] = D_j(\tau_j) - \tau_j \mu_j + \tau_j E[\tilde{d}(t) - v(t)] \qquad \text{(A.1)}$$

In (A.1), the last term corresponds to the expected truncated demand during τ_j periods. The expected inventory level given in (1.16) differs from (A.1) due to this term. By using (A.1), objective function (1.17) can be modified for an exact representation of the considered setting where the truncated demand is lost or handled outside the system.

The expected truncated demand $E[\tilde{d}(t) - v(t)]$ in an arbitrary period is a non-linear function which depends on the net replenishment times of all stages. This makes characterising the last term in (A.1) challenging (at least as challenging as characterising the effectively observed CSL). Besides, if one considers (A.1) in the objective function, finding a solution to the GSM problem would be cumbersome. Consequently, using (1.17) as objective function represents an approximation which simplifies the problem. This approximation will be accurate enough for high CSL targets since demand will be rarely truncated. In this thesis, we consider (1.17) as objective function by referring to the original GSM.

CHAPTER 4: OPTIMISING REORDER INTERVALS AND ORDER-UP-TO LEVELS

As presented in our literature review in Chapter 2, all models in the GSM literature consider the reorder intervals of stages in the supply chain as given input parameters. In this chapter, we study the problem of simultaneously optimising the reorder intervals and order-up-to levels in general acyclic multi-echelon systems facing stochastic demand. Finding an optimal policy for this problem would be extremely difficult. Indeed, the structure of the optimal policy is unknown even for two-echelon distribution systems under periodic-review (Doğru et al., 2009). We build on the Power-of-Two (PO2) and the GSM research to find a reasonable solution to this problem. In order to deal with demand variations, we use the original assumptions of the GSM that are the guaranteed service times and bounded demand assumptions. Besides, we assume that each stage of the supply chain operates with a periodic-review, order-up-to (R, S) policy with stationary nested PO2 reorder intervals. Under this setting, demand bounds are specified using the findings presented in Chapter 3 (Section 3.3).

This chapter has several contributions. First, we propose a deterministic optimisation model for general multi-echelon systems to determine the optimal parameters R and S as well as the corresponding service times. This leads to a Non Linear Integer Programming (NLIP) problem with a non-convex and non-concave objective function including rational and square root terms. Second, we propose a Sequential Optimisation Procedure (SOP) to obtain near optimal solutions with reasonable computational time. We measure the performance of this procedure on randomly generated instances pertaining to two supply chain structures, a five-stage serial and a five-echelon general acyclic system. Third, by defining reasonable bounds for the decision variables of the NLIP model, we propose an Improved Direct (ID) approach.

This chapter is organised as follows. Section 4.1 reviews the literature that is complementary to Chapter 2. In Section 4.2, we develop the NLIP. We then present the SOP in Section 4.3. Section 4.4 establishes the bounds for the decision variables of the NLIP and hence, proposes the ID approach. Numerical analysis on the SOP and the ID approach for

serial and general acyclic multi-echelon systems are summarised in Section 4.5. Finally, Section 4.6 draws some conclusions and suggests potential future research directions.

4.1 Related Literature

This work is built upon two research streams which are the problem of safety stock optimisation and the problem of determining optimal reorder intervals in multi-echelon systems. We refer the reader to Chapter 2 for a literature review of the multi-echelon safety stock optimisation problem. In this section, we provide a brief literature review of the problem of determining optimal reorder intervals in multi-echelon systems.

The optimal reorder intervals that minimise the total cost (including fixed ordering costs) in a multi-echelon system are often impractical to implement since they can take any positive real value. A more realistic problem would then be obtained by assuming the reorder intervals as multiples of a base planning period (e.g., a day, a week or a month). This refers to a replenishment policy in which the reorder interval of each stage is an integer multiple of the base planning period. This policy is known as the integer-ratio policy. Most existing papers consider a subset of integer-ratio policies, the so-called PO2 policy. Under a PO2 policy, reorder intervals are power-of-two multiples of the base planning period. The practical advantages of using a PO2 policy in a multi-echelon inventory system is discussed in Muckstadt and Roundy (1993) and Muckstadt and Sapra (2010). They show that PO2 policies may considerably reduce time and resource consumption in scheduling operations. Furthermore, such policies are efficient for multi-echelon systems facing deterministic demand. With available algorithms, an optimal PO2 solution can be found easily and the solution is guaranteed to be within 6% of optimality if the base planning period is fixed (see Roundy, 1985a) and 2% of optimality if the base planning period is treated as a variable (see Roundy, 1985b; Roundy, 1986).

For serial systems facing stochastic demand, it is observed that the integer-ratio policies obtained by solving the deterministic counterpart of the problem can be an effective heuristic approach. For instance, numerical studies in Chen and Zheng (1998) and Shang (2008) show that the average performance of this deterministic approach is fairly good for

serial systems facing stochastic demand. Chu and Shen (2010) study a two-echelon distribution system facing stochastic demand and having target service levels. They develop a polynomial algorithm to find a PO2 policy whose cost is guaranteed to be no more than 1.26 times the optimal cost. Shang and Zhou (2010) consider a two-echelon distribution system under backordering costs. With a numerical study, they demonstrate that the integer-ratio policy is a good candidate for designing heuristics. However, they show that non-nested PO2 solutions obtained by solving the corresponding deterministic model can perform poorly in some cases. In this chapter, we propose a SOP for general multi-echelon systems facing stochastic demand. Similarly, at the first step of this procedure, we solve the deterministic counterpart of the problem as in the papers mentioned above. We also discuss the conditions under which the quality of our approach deteriorates.

4.2 Optimisation Model

This section presents the optimisation model we propose: Section 4.2.1 introduces our assumptions, Section 4.2.2 shows how to specify the appropriate demand bounds of stages while Section 4.2.3 provides the mathematical programming formulation.

4.2.1 Assumptions

We consider a general multi-echelon system modelled as a network as presented in Section 1.1.1 of Chapter 1. We assume that external demand occurs only at the most downstream stages which we term demand stages. For each demand stage $j \in \mathbf{N_D}$, demand follows a stationary *i.i.d.* process with mean μ_j and standard deviation σ_j per base planning period. For an internal or supply stage $j \in \mathbf{N_I} \cup \mathbf{N_S}$, we can compute the mean demand μ_j per base planning period using (1.1) (see page 12). We consider the case of no risk pooling. Hence, demand variability at stage $j \in \mathbf{N_I} \cup \mathbf{N_S}$ is equal to the sum of demand variability associated with its successors. For $j \in \mathbf{N_I} \cup \mathbf{N_S}$, we define the standard deviation σ_j per base planning period by:

$$\sigma_j = \sum_{k:(j,k)\in\mathbf{A}} \theta_{jk}\sigma_k$$

We assume that demand satisfied from stock is bounded for any long period τ_j and for every stage *j*. As in the original GSM, our model does not address what happens when arrival demand exceeds the specified demand bounds. We assume that demand bounds are defined by the company policy in such a way that the effect of excess demand is tolerated or handled by some extraordinary measures such as subcontracting, overtime production, express expediting etc. The impact of these extraordinary measures on the company is not studied in the GSM neither in our model.

At each stage *j*, we assume a deterministic and constant lead time L_j which corresponds to the duration of process being realised at this stage. Each stage *j* operates with a stationary (R_j, S_j) policy where R_j is the reorder interval and S_j is the base stock level. There is no time delay in ordering. We restrict attention to stationary nested PO2 policies. Thus, the reorder interval R_j can take the following values: $\{1,2,...,2^{l_j}\}$ where l_j is a non-negative integer. Furthermore, since we consider nested policies, the reorder interval of stage $j \in \mathbf{N_I} \cup \mathbf{N_D}$ cannot be greater than the reorder intervals of its upstream stages. We note that the reorder epochs are offset to allow each stage to replenish from its immediate upstream stages at the exact moment an order arrives at the upstream stages and equidistant times of length R_j thereafter.

As in the original GSM, we assume that each stage *j* promises a unique guaranteed outbound service time s_j^{out} to its customers. Demand stages should ensure the maximum service time s_j^{client} tolerated by the final customer. The inbound service time s_j^{in} define the time for stage *j* to get all the inputs from stage $i : (i, j) \in \mathbf{A}$ to start the process. We note that s_j^{out}, s_j^{in} are the decision variables of our optimisation problem. These decision variables serve to determine the safety stock level and the order-up-to level S_j at each stage *j*. We assume that the outbound and the inbound service times are integer multiples of the base planning period likewise the reorder interval. For the sake of simplicity, we will further consider the base planning period as one unit of time and the decision variables as positive integers.

Two types of cost are considered in our model: the fixed ordering and the holding cost. Let A_j be the fixed ordering cost of stage *j* and χ be the number of base planning

periods per year. The Annual Fixed Ordering Cost (AFOC) is calculated similarly to Economic Order Quantity model:

$$\text{AFOC} = \sum_{j \in \mathbf{N}} \frac{A_j}{R_j} \chi$$

The annual holding cost formulation is based on the approximation of Hadley and Whitin (1963) which is the sum of cycle stock and safety stock costs. In order to compute the cycle stock costs, we use the echelon stock approach. Under periodic-review, the on-hand stock evolutions for installation stocks are not of the saw-tooth form at supply and internal stages (see Figure 3.5, Chapter 3). However, the on-hand stock evolutions for echelon stocks are always of the saw-tooth form at all stages no matter the network topology. Thus, it is easier to compute the average echelon stock compared to average on hand stock. Besides, the two approaches yield the same cycle stock costs for the multi-echelon system with nested PO2 policies (Muckstadt and Roundy, 1993). The (annual) per-unit echelon holding cost of stage *j* is denoted by h_j^e (see Section 1.1.4 of Chapter 1 for its definition). The Annual Cycle Stock Cost (ACSC) of the system can be calculated by:

$$\text{ACSC} = \sum_{j \in \mathbf{N}} \frac{1}{2} \mu_j h_j^e R_j$$

The Annual Safety Stock Cost (ASSC) of stage *j* is the product of the (annual) per-unit holding cost h_j and the safety stock level SS_j of stage *j*:

$$\text{ASSC} = \sum_{j \in \mathbf{N}} h_j SS_j$$

In our mathematical model, the cost of pipeline stock is ignored since it depends only on input parameters and does not affect the optimisation. However, this is not to say that the pipeline stock is not a significant part of the inventory in a supply chain. Therefore, the annual cost function that we aim to minimise is the sum of the AFOC, the ACSC and the ASSC.

4.2.2 Demand Bound Functions

The structure of demand bound functions proposed in this section is similar to the one of Section 3.3.1 (Chapter 3). In this section, we introduce some additional notations for ease of exposition while presenting the NLIP model for general acyclic multi-echelon systems.

The existence of guaranteed service times assumption implies that if a stage *j* faces a demand $d_j(t)$ at time *t*, the demand within the demand bounds is fully satisfied with 100% service at time $t + s_j^{out}$. We assume that a replenishment is available to serve demand in its period of arrival. Let consider the replenishment mechanism at an internal or supply stage $j \in \mathbf{N_I} \cup \mathbf{N_S}$. Without loss of generality, stage *j* places orders at times mR_j where *m* is a non-negative integer. Let $t = \varpi + n.R_j$ for $\varpi \in \{1,2,..,R_j\}$. Stage *j* places an order for $d_j(t)$ at time $t + R_j - \varpi$ and the order corresponding to this demand is received at time $t + s_j^{in} + L_j + R_j - \varpi$. In the worst case, $\varpi = 1$ and the reception occurs at time $t + s_j^{in} + L_j + R_j - 1$. If a demand is served first and the replenishment corresponding to this demand occurs at a subsequent period, stage *j* has to store the inventory that would satisfy the demand within the guaranteed service time. That is, if $s_j^{in} + L_j + R_j - 1 > s_j^{out}$, stage *j* should have the amount of inventory to cover the demand over an interval of length $\tau_j = s_j^{in} + L_j + R_j - 1 - s_j^{out}$, that is called the net replenishment time of stage $j \in \mathbf{N_I} \cup \mathbf{N_S}$.

We assume that the external demand occurs continuously over the base planning period. Thus, taking into account an additional increase of the net replenishment time by the base planning period, the net replenishment time τ_j for a demand stage $j \in \mathbf{N_D}$ is equal to $s_j^{in} + R_j + L_j - s_j^{out}$. As in Graves and Willems (2000) one can set the demand bound function for demand stages as follows:

$$D_j(\tau_j) = \tau_j \mu_j + z_j \sigma_j \sqrt{\tau_j} \qquad \text{for } j \in \mathbf{N_D}$$

where z_j is the safety factor of stage *j* that relates to its non-stock-out probability during τ_j.

For internal and supply stages, the maximum demand which can be observed during the net replenishment time depends on the reorder intervals of their immediate downstream stages. The average size of an order placed by stage $k : (j,k) \in \mathbf{A}$ is $R_k \theta_{jk} \mu_k$. The number of

orders placed by stage k and observed by stage j during the net replenishment time τ_j can be calculated by the floor function $\mathrm{n}(\tau_j, R_k)$:

$$\mathrm{n}(\tau_j, R_k) = \left\lfloor \frac{\tau_j}{R_k} \right\rfloor \qquad \text{for } (j,k) \in \mathbf{A}$$

The average demand requested by stage k and observed by stage j during the net replenishment time τ_j is the product of the number of orders placed by stage k during τ_j and the average size of an order placed by stage k. Since we consider the case of no risk pooling, the maximum demand $D_{jk}(\tau_j, R_k)$ placed by stage k and observed by stage j during τ_j can be calculated by:

$$D_{jk}(\tau_j, R_k) = \mathrm{n}(\tau_j, R_k) R_k \theta_{jk} \mu_k + z_j \theta_{jk} \sigma_k \sqrt{\mathrm{n}(\tau_j, R_k) R_k} \qquad \text{for } (j,k) \in \mathbf{A}$$

To provide a guaranteed service time at stage j, the order-up-to level S_j should be equal to the demand upper bound during its net replenishment time:

$$S_j = D_j(\tau_j) \qquad \text{for } j \in \mathbf{N_D}$$

$$S_j = \sum_{k:(j,k)\in\mathbf{A}} D_{jk}(\tau_j, R_k) \qquad \text{for } j \in \mathbf{N_I} \cup \mathbf{N_S}$$

Thus, the safety stock level SS_j at stage j becomes:

$$SS_j = z_j \sigma_j \sqrt{\tau_j} \qquad \text{for } j \in \mathbf{N_D}$$

$$SS_j = z_j \sum_{k:(j,k)\in A} \theta_{jk} \sigma_k \sqrt{\mathrm{n}(\tau_j, R_k) R_k} \qquad \text{for } j \in \mathbf{N_I} \cup \mathbf{N_S}$$

In the mathematical model, we will represent the safety stock function of internal or supply stages without referring to the floor function. Let η_{jk} be the decision variables of the mathematical model representing the floor function value $\mathrm{n}(\tau_j, R_k)$. The variables η_{jk} should verify the following constraints:

$$\tau_j - \eta_{jk} R_k \geq 0 \qquad \forall (j,k) \in \mathbf{A}$$

$$\tau_j - \eta_{jk} R_k < R_k \qquad \forall (j,k) \in \mathbf{A}$$

$$\eta_{jk} \geq 0 \quad \text{and integer} \quad \forall (j,k) \in \mathbf{A}$$

So considered, the term in the square root can be replaced by $\eta_{jk} R_k$. The safety stock level for a stage $j \in \mathbf{N_I} \cup \mathbf{N_S}$ is then equal to:

$$SS_j = z_j \sum_{k:(j,k)\in\mathbf{A}} \theta_{jk} \sigma_k \sqrt{\eta_{jk} R_k} \quad \text{for } j \in \mathbf{N_I} \cup \mathbf{N_S}$$

where η_{jk} verifies the constraints above.

4.2.3 Mathematical Programming Formulation

The problem *P*3 of finding the optimal PO2 reorder intervals and guaranteed service times in order to minimise the total annual cost of the multi-echelon system can be formulated as follows:

$$P3: \min \sum_{j\in\mathbf{N}} \left(\frac{A_j}{R_j} \chi + \frac{1}{2} u_j h_j^e R_j \right) + \sum_{j\in\mathbf{N_I}\cup\mathbf{N_S}} \sum_{k:(j,k)\in\mathbf{A}} h_j z_j \theta_{jk} \sigma_k \sqrt{\eta_{jk} R_k} + \sum_{j\in\mathbf{N_D}} h_j z_j \sigma_j \sqrt{\tau_j} \tag{4.1}$$

s.t. $R_j = 2^{l_j} \qquad \forall j \in \mathbf{N}$ (4.2)

$R_i \geq R_j \qquad \forall (i,j) \in \mathbf{A}$ (4.3)

$\tau_j = s_j^{in} + L_j + R_j - s_j^{out} \qquad \forall j \in \mathbf{N_D}$ (4.4)

$\tau_j = s_j^{in} + L_j + R_j - 1 - s_j^{out} \qquad \forall j \in \mathbf{N_I} \cup \mathbf{N_S}$ (4.5)

$\tau_j - \eta_{jk} R_k \geq 0 \qquad \forall (j,k) \in \mathbf{A}$ (4.6)

$\tau_j - \eta_{jk} R_k < R_k \qquad \forall (j,k) \in \mathbf{A}$ (4.7)

$s_j^{in} \geq s_i^{out} \qquad \forall (i,j) \in \mathbf{A}$ (4.8)

$s_j^{out} \leq s_j^{client} \qquad \forall j \in \mathbf{N_D}$ (4.9)

$l_j, s_j^{in}, s_j^{out}, \tau_j \geq 0 \quad \text{and integer} \qquad \forall j \in \mathbf{N}$ (4.10)

$$R_j \geq 1 \qquad \text{and integer} \qquad \forall j \in \mathbf{N} \tag{4.11}$$

$$\eta_{jk} \geq 0 \qquad \text{and integer} \qquad \forall (j,k) \in \mathbf{A} \tag{4.12}$$

The decision variables of this problem are: the reorder intervals (R_j), the integer variables representing the PO2 values (l_j), the net replenishment times (τ_j), the outbound service times (s_j^{out}), the inbound service times (s_j^{in}) and the number of orders placed by stage k to stage j during τ_j (η_{jk}) for each $(j,k) \in \mathbf{A}$.

The problem $P3$ minimises the total cost function (4.1). Constraint (4.2) restricts the reorder intervals to PO2 solutions. Constraint (4.3) is necessary to ensure nestedness. Constraints (4.4) and (4.5) give the net replenishment times of stages. The nonlinear constraints (4.6) and (4.7) determine the number of orders placed by an internal stage during the net replenishment time of its immediate upstream stage. Constraint (4.8) ensures that the outbound service time of a stage's immediate upstream stage is no greater than its inbound service time. Constraint (4.9) ensures that the demand stages satisfy their service guarantee. With constraints (4.10)-(4.12) decision variables are forced to be positive integers.

The problem $P3$ is a NLIP problem with a neither convex nor concave objective function on the feasible region (see the Appendix of this chapter) including rational and square root terms.

After having solved the problem $P3$, the optimal order-up-to levels S_j^* of stages can be obtained by:

$$S_j^* = D_j(\tau_j^*) \qquad \text{for } j \in \mathbf{N_D}$$

$$S_j^* = \sum_{k:(j,k)\in\mathbf{A}} D_{jk}(\tau_j^*, R_k^*) \qquad \text{for } j \in \mathbf{N_I} \cup \mathbf{N_S}$$

where τ_j^* and R_j^* are the optimal solutions of the problem $P3$.

4.3 Sequential Optimisation Procedure

For large multi-echelon systems the problem $P3$ becomes computationally intractable with direct solution approaches because of the combinatorial nature of the problem and

nonlinear non-convex terms. We thus propose a Sequential Optimisation Procedure (SOP) to obtain near optimal solutions with reasonable computational time. Our method consists of two optimisation procedures. First, we determine the convenient reorder intervals using available optimisation models for nested PO2 policies with deterministic demand. Second, we obtain convenient order-up-to levels, guaranteed service times and safety stock placements using the results of the first procedure as input parameters.

The first optimisation procedure aims at determining a nested PO2 solution to the deterministic counterpart of this problem. Therefore, we first consider the problem *P*4:

$$P4: \min \sum_{j \in \mathbf{N}} \frac{A_j}{R_j} \chi + \frac{1}{2} u_j h_j^e R_j \tag{4.13}$$

$$\text{s.t. } R_j = 2^{l_j} \qquad \forall j \in \mathbf{N} \tag{4.14}$$

$$R_i \geq R_j \qquad \forall (i, j) \in \mathbf{A} \tag{4.15}$$

$$l_j \geq 0 \text{ and integer} \quad \forall j \in \mathbf{N} \tag{4.16}$$

$$R_j \geq 1 \text{ and integer} \quad \forall j \in \mathbf{N} \tag{4.17}$$

The problem *P*4 is studied in the literature for general acyclic multi-echelon systems. To find an optimal solution to this problem, one can use the polynomial time algorithm presented by Maxwell and Muckstadt (1985) and Muckstadt and Roundy (1993).

Let R_j^{seq} be the reorder interval of stage *j* obtained by solving the problem *P*4. By considering reorder intervals as input parameters, reorder interval of stage *j*, R_j^{seq} can be aggregated into its lead time L_j. Hence, lead times of stage *j* can be replaced by L'_j where:

$$L'_j = L_j + R_j^{seq} \qquad \text{for } j \in \mathbf{N_D}$$

$$L'_j = L_j + R_j^{seq} - 1 \qquad \text{for } j \in \mathbf{N_I} \cup \mathbf{N_S}$$

Therefore, the problem *P*3 can be reduced to the problem *P*5:

$$P5: \min \sum_{j \in \mathbf{N}} C_j\left(s_j^{in}, s_j^{out}\right) \tag{4.18}$$

$$\text{s.t. } s_j^{in} + L_j' - s_j^{out} \geq 0 \qquad \forall j \in \mathbf{N} \tag{4.19}$$

$$s_j^{in} \geq s_i^{out} \qquad \forall (i,j) \in \mathbf{A} \tag{4.20}$$

$$s_j^{out} \leq s_j^{client} \qquad \forall j \in \mathbf{N_D} \tag{4.21}$$

$$s_j^{in}, s_j^{out} \geq 0 \text{ and integer} \quad \forall j \in \mathbf{N} \tag{4.22}$$

where;

$$C_j\left(s_j^{in}, s_j^{out}\right) = h_j z_j \sigma_j \sqrt{s_j^{in} + L_j' - s_j^{out}} \qquad \text{for } j \in \mathbf{N_D}$$

$$C_j\left(s_j^{in}, s_j^{out}\right) = \sum_{k:(j,k)\in\mathbf{A}} h_j z_j \theta_{jk} \sigma_k \sqrt{\left\lfloor \frac{s_j^{in} + L_j' - s_j^{out}}{R_k^{seq}} \right\rfloor R_k^{seq}} \qquad \text{for } j \in \mathbf{N_I} \cup \mathbf{N_S}$$

The second procedure aims at finding an optimal solution to the problem *P*5. The problem *P*5 is a GSM with a non-continuous objective function. The cost function of stages only depends on its own service times and is increasing in s_j^{in} and decreasing in s_j^{out}. Therefore, considering the multi-echelon system structure, generic solution techniques developed by Graves and Willems (2000), Humair and Willems (2006) and Humair and Willems (2011) can be used to solve this problem to optimality. In fact, for these techniques there are no structural limitations on $C_j\left(s_j^{in}, s_j^{out}\right)$ as long as the cost function of stage *j* only depends on s_j^{in} and s_j^{out}.

By solving the problem *P*5, we obtain the best service times for the multi-echelon system given reorder intervals R_j^{seq}. As presented in Section 4.2.3, we can deduce the safety stock and order-up-to levels corresponding to this solution. A feasible solution for the problem *P*3 is then obtained by combining the solutions found for problems *P*4 and *P*5.

4.4 Improved Direct Approach

A direct approach to obtain a global optimal solution for the problem *P*4 is to solve it by using a global optimiser such as BARON with 0% optimality margin. BARON provides global optima for this problem if finite lower and upper bounds on the decision variables are

properly specified. When the default decision variable bounds are too large, this approach requires significant computational time (see Section 4.5). Otherwise, if these bounds are too tight, global optima may not be obtained. We improve this Default Direct (DD) approach: first, we establish the solution obtained by the SOP as an initial solution. Second, we develop appropriate decision variable bounds using the solution obtained by the SOP.

In what follows, we show how to establish the decision variable bounds in order to develop an Improved Direct (ID) approach. By solving the problem *P*4, we obtain for each stage *j*, the reorder interval R_j^{seq} that optimises the convex part of the cost function including the annual fixed ordering cost and the annual cycle stock cost. If the optimal reorder interval of a supply stage $j \in \mathbf{N_S}$ is greater than R_j^{seq} found by the sequential optimal solution, the cost of the convex part increases. However, this also increases the net replenishment time of stage $j \in \mathbf{N_S}$ and hence, its annual safety stock cost. Therefore, it is not beneficial for a supply stage to set a reorder interval greater than its R_j^{seq}. Besides, since we only consider nested policies, the reorder interval of a non-supply stage must be smaller than or equal to the maximum reorder interval of the supply stages. Hence, we can establish the upper bounds for all reorder intervals by:

$$R_j \le \max\left\{R_j^{seq} \mid j : j \in \mathbf{N_S}\right\} \quad \text{for } j \in \mathbf{N} \tag{4.23}$$

Similarly, the upper bound for the integer decision variable l_j becomes:

$$l_j \le \max\left\{l_j^{seq} \mid j : j \in \mathbf{N_S}\right\} \quad \text{for } j \in \mathbf{N} \tag{4.24}$$

Since an upper bound can be defined for reorder intervals we can deduce upper bounds for service times as well. We can define the maximum replenishment time M_j by:

$$M_j = L_j - 1 + \max\left\{R_j^{seq} \mid j : j \in \mathbf{N_S}\right\} \quad \text{for } j \in \mathbf{N_S}$$

$$M_j = L_j - 1 + \max\left\{R_j^{seq} \mid j : j \in \mathbf{N_S}\right\} + \max\left\{M_i \middle| i : (i, j) \in \mathbf{A}\right\} \quad \text{for } j \in \mathbf{N_I}$$

$$M_j = L_j + \max\left\{R_j^{seq} \mid j : j \in \mathbf{N_S}\right\} + \max\left\{M_i \middle| i : (i, j) \in \mathbf{A}\right\} \quad \text{for } j \in \mathbf{N_D}$$

The total cost increases when the inbound service times or the net replenishment times increase and when the outbound service times decrease. Hence, as in the original GSM, there always exists an optimal solution for the problem *P*3 such that all inbound service times of the supply stages are equal to 0 and the inbound service time of each non-supply stage is equal to the maximum service time of its upstream stages (see Lesnaia, 2004). Therefore, we can establish upper bounds for the inbound and outbound service times as follows:

$$s_j^{out} \le M_j \qquad \text{for } j \in \mathbf{N} \tag{4.25}$$

$$s_j^{in} \le \max\left\{M_i \middle| i : (i,j) \in \mathbf{A}\right\} \qquad \text{for } j \in \mathbf{N_I} \cup \mathbf{N_D} \tag{4.26}$$

$$s_j^{in} \le 0 \qquad \text{for } j \in \mathbf{N_S} \tag{4.27}$$

Hence, upper bounds for τ_j and η_{jk} become:

$$\tau_j \le M_j \qquad \text{for } j \in \mathbf{N} \tag{4.28}$$

$$\eta_{jk} \le M_j \qquad \text{for } (j,k) \in \mathbf{A} \tag{4.29}$$

Besides, the natural lower bounds for these decision variables are given by constraints (4.10)-(4.12). The ID approach is then obtained by setting the bounds (4.23)-(4.29) to the decision variables and by considering the sequential optimal solution as an initial solution.

4.5 Numerical Analysis

In this section, computational experiments are carried to test the relevancy of the SOP and the direct approaches. Randomly generated five-echelon serial and five-echelon general acyclic multi-echelon systems are used. Although the serial system considered has the same structure as that of Chapter 3, we consider in this section different data concerning lead times, holding and ordering costs in order to conduct an enhanced analysis by considering a much richer data set. Besides, since the model is developed for general acyclic structures, we test the different solution procedures on a real-world general acyclic supply chain structure. These data sets are generated in order to conduct a detailed comparison between the different

solution procedures proposed in terms of solution quality and computational time in different supply chain structures.

This section is organised as follows. Section 4.5.1 presents the data generation procedure. Section 4.5.2 provides results on the global optimal reorder intervals based on illustrative examples. Section 4.5.3 discusses the optimality gap of the SOP and identifies conditions under which the SOP performs relatively bad. Finally, Section 4.5.4 compares the performance of the SOP and the direct approaches in terms of computational time.

4.5.1 Data Generation

For numerical analysis, we consider a five-echelon serial (Figure 3.2, Chapter 3) and a five-echelon general acyclic system (Figure 4.1). Common parameters for both structures are as follows. The length of the base planning period is a business day and there are 260 business days in a year. The safety factors for all stages are the same and equal to 1.645 (which correspond to a 95% service level). The maximum service times at demand stages are set to 0. The parameter θ_{jk} is set equal to 1 for all stages $(j,k) \in \mathbf{A}$. Besides, we consider different groups as in Bossert and Willems (2007) based on the ordering cost ratios (i.e. the ratio A_j / h_j for stage j). For the serial five-echelon structure, we define three profiles for each group (see Table 4.1). For each echelon of the five-echelon general acyclic system, we specify intervals in which the ordering cost ratios are generated randomly if the ratio is not set to 0 (see Table 4.2).

For the five-echelon serial supply chain system, mean and standard deviation of daily demand at the demand stage (Stage 5) are respectively $\mu = 150$ and $\sigma = 45$. The (annual) per-unit holding cost for Stage 1 is generated randomly in U[0, 20]. Then, the (annual) per-unit holding costs of other stages are obtained by adding a random number in U[0, 20] to the per-unit holding cost of its upstream stage. The lead time value of each stage is an integer, generated randomly in U[1, 20]. Following these rules, 15 instances that comprise the lead time and the per-unit holding cost data are obtained. The 15 instances (Table 4.3) permuted with the 21 ordering cost profiles (Table 4.1) generate the 315 test problems considered.

The five-echelon general acyclic system corresponds to the real-world supply chain presented by Willems (2008). For this system, we use data provided by Willems (2008) that includes the lead times (the average values are considered), the stage costs (holding cost rate is set to 10%) and mean and standard deviation of demand at demand stages (see Table 4.4). For the ordering costs, we generate 15 instances for each of the 21 ordering cost profiles using intervals reported in Table 4.2. Hence, we obtain 315 test problems.

The SOP and direct approaches are coded in GAMS 23.7 on a VAIO computer with Intel Core i3-2310M processor (2.10 GHz) and 4 GB RAM. BARON (version 9.3.1) is used for the computational experiments. For all test problems, the global optimal (GO) solutions are obtained by the ID approach using the global optimiser BARON. The sequential optimal (SO) solutions are the feasible solutions obtained from the SOP.

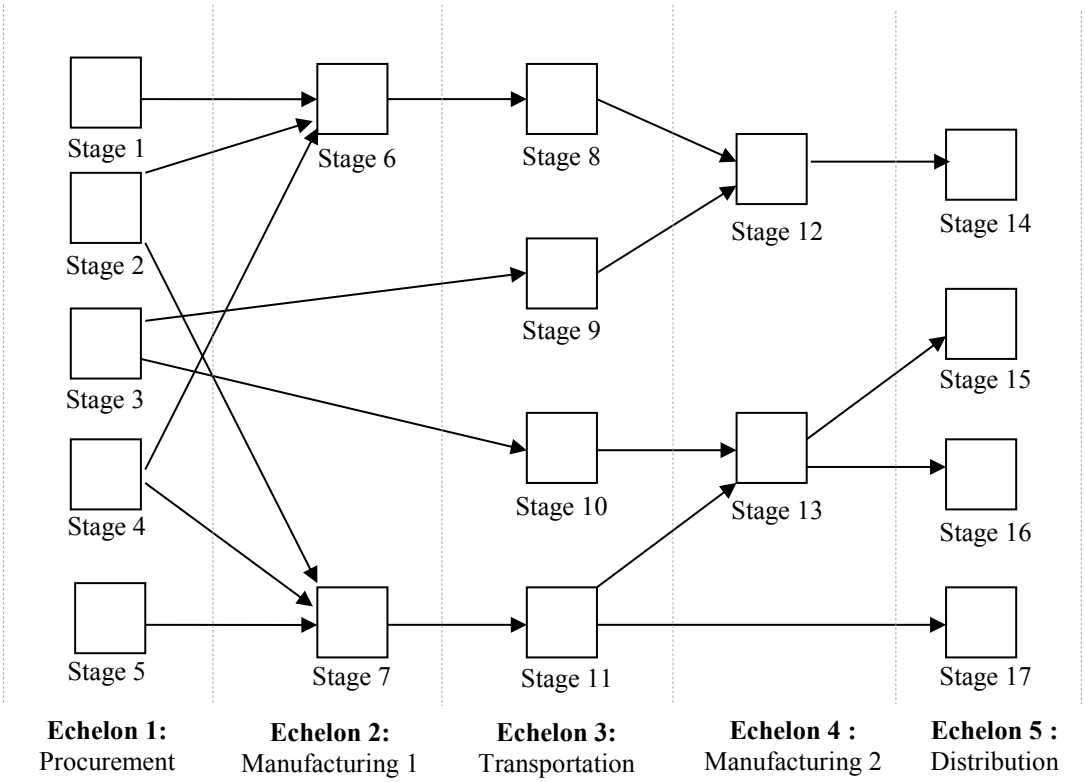

Figure 4.1: The five-echelon general acyclic system

Ordering Cost Profiles	Stages				
	1	*2*	*3*	*4*	*5*
Starting	4	0	0	0	0
	25	0	0	0	0
	80	0	0	0	0
Middle	0	0	4	0	0
	0	0	25	0	0
	0	0	80	0	0
Ending	0	0	0	0	4
	0	0	0	0	25
	0	0	0	0	80
Uniform	2	2	2	2	2
	25	25	25	25	25
	50	50	50	50	50
Increasing	0	2	2	2	25
	0	2	4	25	80
	1	2	16	25	50
Decreasing	25	2	2	2	0
	80	25	4	2	0
	50	25	16	2	1
Random	11	8	9	0	2
	13	25	0	4	12
	60	10	8	24	11

Table 4.1: Ordering cost ratios used for the serial system

Ordering Cost Profiles	Echelons				
	1	*2*	*3*	*4*	*5*
Starting	U[0, 5]	0	0	0	0
	U[15, 25]	0	0	0	0
	U[65, 80]	0	0	0	0
Middle	0	0	U[0, 5]	0	0
	0	0	U[15, 25]	0	0
	0	0	U[65, 80]	0	0
Ending	0	0	0	0	U[0, 5]
	0	0	0	0	U[15, 25]
	0	0	0	0	U[65, 80]
Uniform	U[5, 10]	U[5, 10]	U[5, 10]	U[5, 10]	U[5, 10]
	U[20, 25]	U[20, 25]	U[20, 25]	U[20, 25]	U[20, 25]
	U[45, 50]	U[45, 50]	U[45, 50]	U[45, 50]	U[45, 50]
Increasing	U[0, 5]	U[5, 10]	U[10, 15]	U[15, 20]	U[20, 25]
	U[0, 10]	U[10, 20]	U[20, 30]	U[30, 40]	U[40, 50]
	U[0, 15]	U[15, 30]	U[30, 45]	U[45, 60]	U[60, 75]
Decreasing	U[20, 25]	U[15, 20]	U[10, 15]	U[5, 10]	U[0, 5]
	U[40, 50]	U[30, 40]	U[20, 30]	U[10, 20]	U[0, 10]
	U[60, 75]	U[45, 60]	U[30, 45]	U[15, 30]	U[0, 15]
Random	U[0, 15]	U[0, 15]	U[0, 15]	U[0, 15]	U[0, 15]
	U[0, 30]	U[0, 30]	U[0, 30]	U[0, 30]	U[0, 30]
	U[0, 50]	U[0, 50]	U[0, 50]	U[0, 50]	U[0, 50]

Table 4.2: Ordering cost ratio intervals used for the general acyclic system

Instance	Stage j	L_j	h_j(\$)	**Instance**	Stage j	L_j	h_j(\$)	**Instance**	Stage j	L_j	h_j(\$)
	1	10	15.9		1	16	15.6		1	17	5.3
	2	14	32.3		2	8	22.7		2	15	16.5
1	3	2	36.4	6	3	1	25.1	11	3	6	23.0
	4	8	55.8		4	15	30.4		4	6	37.1
	5	11	61.1		5	16	43.7		5	6	49.3
	1	16	13.2		1	20	9.3		1	8	12.9
	2	5	28.4		2	13	12.3		2	11	20.5
2	3	12	36.9	7	3	15	29.5	12	3	19	31.7
	4	20	53.4		4	4	42.4		4	1	47.8
	5	13	58.5		5	20	56.7		5	12	49.1
	1	15	10.8		1	10	14.6		1	5	8.9
	2	6	15.6		2	9	19.6		2	11	12.9
3	3	1	18.7	8	3	9	29.8	13	3	4	13.0
	4	12	20.8		4	8	36.4		4	3	21.9
	5	9	24.0		5	8	39.7		5	10	35.3
	1	4	14.8		1	3	15.3		1	16	7.0
	2	18	31.6		2	3	21.4		2	14	19.9
4	3	18	35.7	9	3	12	24.5	14	3	19	28.4
	4	1	43.2		4	2	36.9		4	11	36.9
	5	9	59.6		5	2	43.8		5	13	47.8
	1	20	10.1		1	13	17.9		1	12	10.2
	2	19	13.9		2	19	33.2		2	13	26.0
5	3	5	25.7	10	3	8	45.1	15	3	4	35.5
	4	6	26.0		4	11	50.8		4	5	53.5
	5	16	31.9		5	8	68.6		5	14	69.0

Table 4.3: Lead time and per-unit holding cost data of the serial test problems

Stage j	Lead Time L_j	Holding Cost h_j (\$)	μ_j (External Demand)	σ_j (External Demand)	s_j^{client} (Demand stages)	α_j (Demand stages)
1	45	110	-	-	-	-
2	38	60	-	-	-	-
3	31	150	-	-	-	-
4	54	40	-	-	-	-
5	26	110	-	-	-	-
6	12	240	-	-	-	-
7	10	235	-	-	-	-
8	2	240	-	-	-	-
9	2	150.3	-	-	-	-
10	2	150.2	-	-	-	-
11	2	235	-	-	-	-
12	10	395.3	-	-	-	-
13	10	391.2	-	-	-	-
14	2	410.3	126	132.3	0	95%
15	5	416.2	57	51.3	0	95%
16	5	424.7	46	45	0	95%
17	1	275	70	35	0	95%

Table 4.4: Data provided by Willems (2008) for the general acyclic system

4.5.2 Results

We observe that when the SO and GO solutions are not the same, the GO solution may lead to smaller reorder intervals than the SO solution. In this case, the sum of total annual fixed ordering and annual cycle stock costs ($AFOC_j$ +$ACSC_j$ of stages $j \in \mathbf{N}$) increases in comparison with the SO solution. Besides, this may also increase the annual safety stock cost $ASSC_i$ at the upstream stage(s) $i:(i,j) \in \mathbf{A}$ since stage j will order more frequently during the net replenishment time of stage i. On the other hand, this leads to a potential reduction of the annual safety stock cost at stage j or at the one(s) of its downstream stage(s). The additive effect of these deviations may reduce the total cost. To illustrate this result, we provide in Table 4.5 the SO and GO solutions for a five-echelon serial test problem (obtained by permuting instance 14 with the second decreasing ordering cost profile). For this example, the total cost of the SO and GO solutions are respectively \$90,100 to \$89,208 which represents a relative gap of 1%.

Another interesting result is the reduction of the total cost when the reorder interval of a non-supply stage j is increased over its SO reorder interval. This action may increase the sum $AFOC_j$ +$ACSC_j$ for stage j. However, it may decrease $ASSC_i$ at the upstream stage(s) $i:(i,j)\in \mathbf{A}$. This is due to a better order coordination between customer-supplier stages. Besides, this may also reduce the safety stock cost at other stage(s) sharing a same supplier with stage j since each supplier quotes a unique service time for all of its customers. To illustrate this result we provide in Table 4.6 the GO and SO solutions for a general acyclic test problem pertaining to the first decreasing ordering cost profile. For this example, the total SO and GO costs are respectively \$3,180,765 to \$3,141,906 and this represents a relative gap of 1.24%.

Stage	**The SO Solution**				**The GO Solution**			
j	R_j (days)	s_j^{out} (days)	$ASSC_j$ (\$)	$AFOC_j$+ $ACSC_j$ (\$)	R_j (days)	s_j^{out} (days)	$ASSC_j$ (\$)	$AFOC_j$+ $ACSC_j$ (\$)
1	16	0	2,061	17,451	16	0	2,524	17,451
2	16	22	-	23,508	8	14	-	23,850
3	8	45	-	8,829	8	37	-	8,829
4	4	59	-	7,326	4	51	-	7,326
5	1	0	30,109	817	1	0	28,411	817
Total			32,169	57,931			30,935	58,273

Table 4.5: Solutions obtained for a serial test problem

Concerning the reorder intervals obtained for different groups of ordering cost ratio profiles, ending, uniform and increasing groups lead to the same reorder interval among all stages since we only consider nested policies. In this case, decreasing the reorder intervals of all stages together may improve the SO solution. In the general acyclic structure, this may imply high cost deviations and may significantly reduce the total cost. For decreasing and random groups, the total cost of the system may be reduced by increasing or decreasing the reorder interval of the SO solution for one or several stages. For the starting group the SO solution may be improved by decreasing the reorder interval at supply stages. Similarly for the middle group, a better solution than the SO solution may be found by decreasing the reorder intervals at the first three upstream echelons. However, for this group, the

improvement of the total cost function is restrictive and it usually implies small cost deviations.

Stage	The SO Solution				The GO Solution			
j	R_j (days)	s_j^{out} (days)	$ASSC_j$ ($)	$AFOC_j$+ $ACSC_j$ ($)	R_j (days)	s_j^{out} (days)	$ASSC_j$ ($)	$AFOC_j$+ $ACSC_j$ ($)
1	16	45	-	152,224	16	45	-	152,224
2	16	38	-	163,693	16	38	-	163,693
3	16	31	-	329,261	16	31	-	329,261
4	16	38	69,373	109,106	16	38	69,373	109,106
5	16	26	-	188,098	16	26	-	188,098
6	16	57	-	103,575	16	57	-	103,575
7	16	48	-	95,483	16	48	-	95,483
8	16	59	-	45,912	16	59	-	45,912
9	16	48	-	32,549	16	48	-	32,549
10	16	48	-	36,358	16	48	-	36,358
11	16	58	-	54,822	16	50	-	54,822
12	16	69	-	46,135	16	69	-	46,135
13	16	80	-	38,300	16	60	-	38,300
14	16	0	832,814	35,425	16	0	832,814	35,425
15	16	0	352,945	37,653	16	0	316,074	37,653
16	4	0	296,591	5,146	16	0	282,947	12,858
17	8	0	129,588	25,713	16	0	129,588	29,657
Total			1,681,312	1,499,453			1,630,798	1,511,109

Table 4.6: Solutions obtained for a general acyclic test problem

We note that for the considered general acyclic test problems, the GO reorder intervals usually tend to be the same among all stages. This stems from the benefit obtained due to the order coordination. This benefit is significant since the considered general acyclic system represents high demand variability at demand stages. However, the SOP does not consider demand variability to compute the SO reorder intervals.

4.5.3 Optimality Gap

Table 4.7 summarises the performance of the SOP for both structures and for different groups of ordering cost profiles in terms of optimality gap. Optimality gap is computed by $(Obj_{SO} - Obj_{GO})/Obj_{GO}$ where Obj_{SO} denotes the SO objective value and Obj_{GO} the GO objective value. The average (*Avr.*) and the maximum (*Max.*) gaps are reported for each

ordering cost profile that comprises 45 test problems. The last column (titled "Occ.") reports the number of occurrences where the SO solution is different from the GO solution.

For the serial structure the SO solution is usually equal to the GO solution. Decreasing and random profiles reveal a relatively inferior performance. The SOP is able to obtain near optimal solutions of about 0.01% optimality gap on average with an observed worst-case of 1.23%.

Ordering Cost Profiles	**Serial System**			**General Acyclic System**		
	Avr. Gap	*Max. Gap*	*Occ.*	*Avr. Gap*	*Max. Gap*	*Occ.*
Starting	0.00%	0.00%	0	0.29%	1.27%	31
Middle	0.00%	0.00%	0	0.01%	0.53%	2
Ending	0.00%	0.00%	0	0.60%	3.63%	17
Uniform	0.00%	0.21%	1	0.39%	2.58%	13
Increasing	0.00%	0.00%	0	0.41%	2.22%	15
Decreasing	0.05%	1.00%	5	1.07%	4.87%	27
Random	0.03%	1.23%	1	0.45%	4.61%	14

Table 4.7: Optimality gap results

In the general acyclic structure, similarly to the serial one, the gap is important for decreasing and random groups. Besides, the gap of ending and increasing groups is higher compared to the serial structure. However, the SO solutions are still near optimal with 0.46% optimality gap on average and with an observed worst case of 4.87% for the considered structure.

The demand variability considered in the general acyclic system lies between 0.50 and 1.05. In order to investigate the impact of demand variability on the SOP performance, we perform a second set of experiments for the serial structure. When we increase the coefficient of variation from 0.30 to 1 for this system, the performance of SO solutions deteriorates. Particularly, the performance of the starting, uniform, decreasing and random profiles get worst. The observed worst case still belongs to the random profile with 3.41% optimality gap. However, the SO solutions still represent an average optimality gap of 0.15% for all groups of ordering cost profiles (see Table 4.8).

Ordering Cost Profiles	Serial System		
	Avr. Gap	*Max. Gap*	*Occ.*
Starting	0.01%	0.25%	1
Middle	0.00%	0.00%	0
Ending	0.00%	0.00%	0
Uniform	0.47%	3.36%	12
Increasing	0.00%	0.00%	0
Decreasing	0.37%	2.90%	15
Random	0.22%	3.41%	8

Table 4.8: Optimality gap results for the serial system with a coefficient of variation equal to1

4.5.4 Computational Time

For serial test problems, computational time is less than 1, 2, 6 seconds using respectively the SOP, the ID and DD approaches. Therefore, the computational times of these approaches are very short and similar for five-stage serial structure. However, the differences become significant for the general acyclic structure.

We notice that the DD approach requires significant computational time when the complexity of the supply chain network increases. With the ID approach, the computational time may be significantly reduced. For instance, for general acyclic test problems belonging to the first starting profile, the DD approach cannot converge within 18000 seconds whereas the ID approach provides global optima in 2,785 seconds on average. In this case, the average gap between the best feasible solution obtained by the DD approach and the global optima is about 15.31%. Therefore, the ID approach clearly dominates the DD approach.

Besides, we notice that SOP requires significantly shorter computational time than the ID approach for all profiles. The SOP provides near optimal solutions within 13 seconds whereas the ID approach requires 2,951 seconds on average to provide global optima. Table 4.9 reports the running times of the SOP and the ID approach for all groups of ordering cost profiles.

We observe that for the starting group, computational time of the ID approach is surprisingly long. Using decision variable bounds presented in Section 4.4, decision variable bounds of non-supply stages remain too large for the starting group and this prevents a fast convergence of BARON to global optima. This also explains the relatively long

computational times for middle and decreasing groups. A converse effect is observed in ending, increasing and uniform groups.

Ordering Cost Profiles	SOP			ID Approach		
	Avr. Time (secs)	*Min. Time (secs)*	*Max. Time (secs)*	*Avr. Time (secs)*	*Min. Time (secs)*	*Max. Time (secs)*
Starting	7.57	3.37	12.43	18684.17	161.60	69699.39
Middle	6.90	4.36	12.18	439.69	102.68	1085.88
Ending	4.84	2.27	8.96	122.90	2.61	489.28
Uniform	3.77	2.62	8.36	214.18	13.90	1472.73
Increasing	2.99	1.91	4.15	105.53	7.41	753.98
Decreasing	5.25	2.85	11.59	853.43	26.60	4600.47
Random	4.75	1.81	9.33	243.90	9.27	3041.4

Table 4.9: Running times of the SOP and ID approaches for the general acyclic test problems

4.6 Conclusion

In this chapter, we have presented a NLIP model that determines nested PO2 reorder intervals and order-up-to levels in a multi-echelon inventory system. The GSM approach is used to model the multi-echelon system facing stochastic demand. Our computational studies demonstrate that the performance of the solution procedure may deteriorate when demand variability and the complexity of the supply chain network increases. However, for a five-echelon general acyclic multi-echelon system with 17 stages and 18 arcs facing high demand variability, the SOP provides near optimal solutions of about 0.46% optimality gap on average within 13 seconds. Besides, we also propose an improved direct approach to reduce the computational time when the problem is solved to global optimality using a global optimiser. For test problems for which the global optimiser cannot converge within 18000 seconds, the improved direct approach provides global optima in 2785 seconds on average.

Some additional relevant issues remain for future consideration. The first one is the performance evaluation of the sequential optimisation procedure for more complex and larger multi-echelon systems. For those systems, a faster global optimisation method must be developed in order to realise this analysis. The second issue is the extension of the model to consider non-nested policies which would be more relevant for general supply chain

structures. By considering a non-nested policy, a better solution in terms of total supply chain cost may be obtained. The third issue concerns the relaxation of the model so that stages are allowed to have arbitrary integer reorder intervals. This extension would enable to estimate the cost of the PO2 restrictions. In addition, it seems worthwhile to conduct a comparison of our model with a model that deals with the same problem by employing the SSM approach. However, since the two approaches are based on different settings, how to perform a fair comparison remain as an open question. All these extensions represent challenging future research directions for general multi-echelon systems facing stochastic demand.

Appendix: Neither Convex nor Concave Objective Function

In order to show that objective function (4.1) given in Section 4.2.3 is neither convex nor concave on the feasible region of the problem *P*3, we consider the Hessian matrix of function (4.1) corresponding to the variables R_j where $j \in \mathbf{N_S}$ and to τ_k where $k \in \mathbf{N_D}$ is:

$$\mathbf{H} = \begin{pmatrix} \dfrac{2A_j}{R_j^3} & 0 \\ 0 & -\dfrac{h_k z_k \sigma_k}{4(\tau_k)^{3/2}} \end{pmatrix}$$

We assume that the input parameters for stage j and k are strictly positive. The determinant of matrix $\mathbf{H}$ is then strictly negative when $R_j > 0$ and $\tau_k > 0$. Therefore, function (4.1) is neither convex nor concave with respect to R_j and τ_k on the feasible region of the problem *P*3. Hence, function (4.1) is neither convex nor concave with respect to all of its variables on the feasible region of the problem *P*3.

CHAPTER 5: MITIGATING THE CYCLE-SERVICE-LEVEL DEVIATION IN GUARANTEED-SERVICE SUPPLY CHAINS

Guaranteed-service supply chains represent multi-echelon systems operating under the GSM assumptions. In such systems, demand is assumed to be bounded at each stage of the supply chain. Most studies in the GSM literature specify the demand bounds using a safety factor that relates to a target Cycle-Service-Level (CSL). The effectively observed CSL in guaranteed-service supply chains can be defined as the probability that the safety stocks in the system cover demand variations. We showed in Chapter 3 (Section 3.1) that under the existence of demand bounds, the effectively observed CSL at demand stages may be less than the one used to define the demand bounds. The deviation is due to the fact that the CSL at a demand stage is affected by the demand bounds applied at its upstream stages. This particularly happens when the net replenishment times of upstream-downstream stages are different and the demand bounds in the system are incompatible due to the associated safety factors.

The issue regarding the CSL deviation in guaranteed-service supply chains is not elaborately studied in the literature. In the first part of this chapter, we show how to specify the safety factors to be applied at different stages of the supply chain so that the effectively observed CSL at the most downstream stage achieves the target CSL. To do this, we propose two approaches that mitigate the CSL deviation in guaranteed-service supply chains. The first approach determines a common safety factor value for all stages of the supply chain that enables to achieve the target CSL at the most downstream stage while the second approach adjusts the safety factors according to the per-unit holding costs and the net replenishment times of stages. We focus on serial and assembly systems since we use results presented in Chapter 3 (Section 3.1). The numerical study shows that the first approach outperforms the second one in terms of computational time while the second approach provides better results in terms of solution quality.

In the second part of this chapter, we study the issue of the CSL deviation in a context in which different actors control different parts of the supply chain. Recently, researchers

consider the GSM in decentralised supply chains. However, how to specify the safety factors at different parts of a decentralised supply chain is not discussed in the existing literature. Existing works ignore the conflict that may occur due to the incompatible demand bounds used by different actors of the supply chain. Indeed, if each actor mitigates the CSL deviation independently, the most downstream actor may still face a deviation in its service level due to the demand bounds applied at the upstream actors. We measure the relevant deviation considering five-stage serial and five-echelon assembly systems controlled by two actors.

This chapter is organised as follows. 5.1 presents the mitigation approaches proposed. Section 5.2 presents the CSL deviation issue in decentralised supply chains. Section 5.3 draws conclusions and proposes potential future research directions.

5.1 Mitigation Approaches

In this section, we present two approaches in order to mitigate the CSL deviation in guaranteed-service supply chains by focusing on serial and assembly systems.

In the GSM literature, Minner (2000) is the only one who attempts to present the CSL deviation under the GSM setting and mitigate this deviation by adjusting the safety stock levels of stages. His approach consists of sequentially increasing the safety stock levels (which is equivalent to increase the safety factors). Safety stock adjustments are solely analysed for stages with strictly positive net replenishment times, i.e. for stages that hold safety stocks. Extreme strategies such as increasing only the upstream or only the downstream stock level are evaluated considering two-stage service level interactions. He illustrates these strategies on a three-stage serial system example. He shows that increasing only the downstream safety stock is better in terms of cost compared to increasing only the upstream safety stock. However, this result is limited to a single example. Besides, since he considers two-stage service level interactions, his approach does not enable to achieve the target CSL in our setting when there exist more than two stages holding safety stock in the system. In addition, he uses the formulas presented by Van Houtum et al. (1996) in order to calculate the effectively observed CSL which are shown in Section 3.1.2 of Chapter 3 to be

inappropriate in the setting that we consider where the unbounded demand is truncated and handled outside the normal supply chain.

Our aim is to propose two approaches in order to mitigate the CSL deviation in the GSM setting where the effectively observed CSL is defined as the probability that the safety stocks in the system cover demand variations. The first approach (Section 5.1.1) consists of determining a common safety factor value for all stages of the supply chain that enables to achieve the target CSL at the most downstream stage. In the second approach (Section 5.1.2), safety factors are adjusted according to the per-unit holding costs and the net replenishment times of stages. Hence, each stage may have a different safety factor value after the adjustment of safety factors. Besides, several safety stock placement solutions are evaluated in order to find a solution that minimises the resulting safety stock cost increase. We compare the two approaches in terms of solution quality and computational time using five-stage serial and real-world assembly test problems presented in Section 3.1.3 of Chapter 3 (Section 5.1.3).

5.1.1 Mitigation Approach I

In the GSM literature, most researchers apply a common safety factor value to all stages of the supply chain while specifying demand bounds in serial and assembly systems. As presented in (3.1) (Chapter 3, page 66), this leads to the following demand bound function:

$$D(\tau_j) = \mu\tau_j + z\sigma\sqrt{\tau_j} \text{ for } j \in \{1,2,...,n\}$$

The common safety factor z used for all stages of the supply chain is an input parameter which relates to a target CSL α. As shown in Chapter 3 (Section 3.1), the effectively observed CSL at the most downstream stage may deviate from the target CSL under this setting. In order to remedy this issue, the common safety factor can be adjusted by increasing its value until the target CSL is reached at the most downstream stage. The relevant adjustment of the safety factor implies to increase the safety stock levels of all stages holding safety stocks.

In the first mitigation approach, in accordance with the GSM literature, we propose to use a common safety factor for all stages. We apply the following simple rules to determine the appropriate common safety factor value $\hat{z}$. First, we use the safety factor $z = \Phi^{-1}(\alpha)$ that relates to the target CSL α at all stages of the supply chain. As such, we obtain an initial solution by solving *P*0 (Section 1.2.2, page 30). Then, we measure the effectively observed CSL α' using (3.3) (Section 3.1.1, page 67) under the corresponding demand bounds. If the effectively observed CSL is less than the target one for this initial solution, we increase the common safety factor until the resulting service level deviation is zero or below a certain tolerance level.

We define $z^{\max}$ as an input parameter which is sufficiently large so that the resulting α' is greater than the target one. The appropriate safety factor $\hat{z}$ which mitigates the CSL deviation can be determined by a bisection method:

Step 1. Start with the interval $I = [z, z^{\max}]$

Step 2. Halve the length of *I* so that the common safety factor that ensures the target CSL α is within the new interval.

Step 3. If the length of *I* is not below the pre-specified tolerance level go to Step 2. Otherwise, go to Step 4.

Step 4. Re-calculate the safety stock levels and the total safety stock cost for $\hat{z} \in I$.

The common safety factor $\hat{z}$ that ensures the target CSL at the most downstream stage is within the interval obtained at the final iteration. Here, the error can be no more than the length of the final interval and is restricted considering a certain tolerance level. Under the common safety factor $\hat{z}$ obtained, we recalculate the safety stock levels and the total safety stock cost in the system. The total safety stock cost increases since $\hat{z} > z$. The solution obtained under $\hat{z}$ ensures $|\alpha - \alpha'| \leq \varepsilon$ where $\varepsilon > 0$ represents the tolerance level considered for the CSL deviation.

We note that since we consider a common safety factor $\hat{z}$ for all stages of the supply chain, the solution obtained after mitigation of the CSL deviation leads to the same safety stock placements as the initial solution. That is because, when the safety factor is the same at

all stages, the optimal safety stock placements and net replenishment times become insensitive to the safety factor.

5.1.2 Mitigation Approach II

In the second mitigation approach, we first propose a safety stock adjustment procedure that takes into account the per-unit holding costs and net replenishment times of stages for a given safety stock placement solution. Besides, we evaluate several solutions in order to find a good solution that minimises the total safety stock cost increase resulting from the mitigation of the CSL deviation. Figure 5.1 gives the flowchart diagram of this mitigation approach.

As in the first approach, we first start with determining an initial solution using the safety factor $z = \Phi^{-1}(\alpha)$ for all stages. We measure the effectively observed CSL α' using (3.3) under the corresponding demand bounds and net replenishment times. If the effectively observed CSL is less than the target one for the initial solution, we apply a safety factor adjustment procedure for stages having strictly positive net replenishment times, i.e. for stages holding safety stocks. We denote the set of stages with strictly positive net replenishment times by **T**. We assume that the safety factor of a stage cannot be increased more than a maximum level $z^{\max}$. The safety factor $z^{\max}$ is sufficiently large so that if the safety factor of all stages $\tau_j^* > 0$ equal $z^{\max}$ the resulting α' is greater than the target CSL α. The idea is to increase the safety factor of stages one by one starting from stages with low per-unit holding cost and net replenishment time. Hence, we select stages for which the safety factor to be increased according to their $h_j \tau_j^*$ value. This may limit the resulting safety stock cost increase after the adjustment of safety factors.

Since we increase the safety factor of stages one by one, increasing the safety factor of a stage to its maximum value $z^{\max}$ may not be sufficient to reach the target CSL (especially at the beginning of this procedure). That is, the effectively observed CSL obtained for $\hat{z}_j = z^{\max}$ for $j \in \mathbf{T}$ may be smaller than the target CSL α. Let the effectively observed CSL $\alpha' = \alpha'^{\max}$ for $\hat{z}_j = z^{\max}$. If $\alpha'^{\max}$ is smaller than the target CSL α, i.e. if $\alpha - \alpha'^{\max} > \varepsilon$, we set the safety factor $\hat{z}_j$ to the smallest value that yields the greatest

increase in the effectively observed CSL. In other words, we select the smallest safety factor $\hat{z}_j$ which verifies $|\alpha' - \alpha'^{\max}| \leq \varepsilon$.

If for $\hat{z}_j = z^{\max}$ the effectively observed CSL is greater than the target one, i.e. if $\alpha'^{\max} - \alpha > \varepsilon$, then we determine the safety factor which verifies $|\alpha - \alpha'| \leq \varepsilon$ using a bisection method (as presented in Section 5.1.1). The procedure terminates when $|\alpha - \alpha'| \leq \varepsilon$ under the adjusted safety factors. We note that there is at least one feasible solution that satisfies the termination criteria since $z^{\max}$ is sufficiently large. The relevant safety factor adjustment procedure can be described as follows:

Step 1. Select stage $j \in \mathbf{T}$ such that $h_j \tau_j^*$ is the smallest among all stages in $\mathbf{T}$.

Step 2. Asses the effectively observed CSL $\alpha'^{\max}$ obtained by setting $\hat{z}_j = z^{\max}$.

If $\alpha - \alpha'^{\max} > \varepsilon$ then set $\hat{z}_j$ to the smallest value that verifies $|\alpha' - \alpha'^{\max}| \leq \varepsilon$. Let $\mathbf{T} = \mathbf{T} \setminus \{j\}$ go to Step 1.

Else set $\hat{z}_j$ to the value that verifies $|\alpha - \alpha'| \leq \varepsilon$ using a bisection method and go to Step 3.

Step 3. Re-calculate the safety stock levels and the total cost under the adjusted safety factors. Let the total cost be *Actual Cost*.

After applying the safety factor adjustment procedure, a solution that ensures the target CSL at the most downstream stage is obtained. We denote the total safety stock cost that corresponds to this solution by *MinCost*. After the adjustment of safety factors, each stage may have a different safety factor value. In this case, if the problem *P*0 is re-solved under the adjusted safety factors, the safety stock placements (hence the net replenishment times) obtained may be different from those of the initial solution. If the objective value of *P*0 under the adjusted safety factors is smaller than *MinCost*, it is possible to obtain a better solution than the actual solution by re-adjusting safety factors for this new solution. Hence, we can repeat the safety factor adjustment procedure for this new solution and compare the cost obtained with *MinCost*. If *Actual Cost* is smaller than *MinCost*, we set *MinCost* to *Actual Cost*, we re-solve *P*0 under the adjusted safety factors and then repeat the previous steps. Otherwise, *MinCost* represents the smallest cost obtained and we terminate the mitigation procedure (see Figure 5.1).

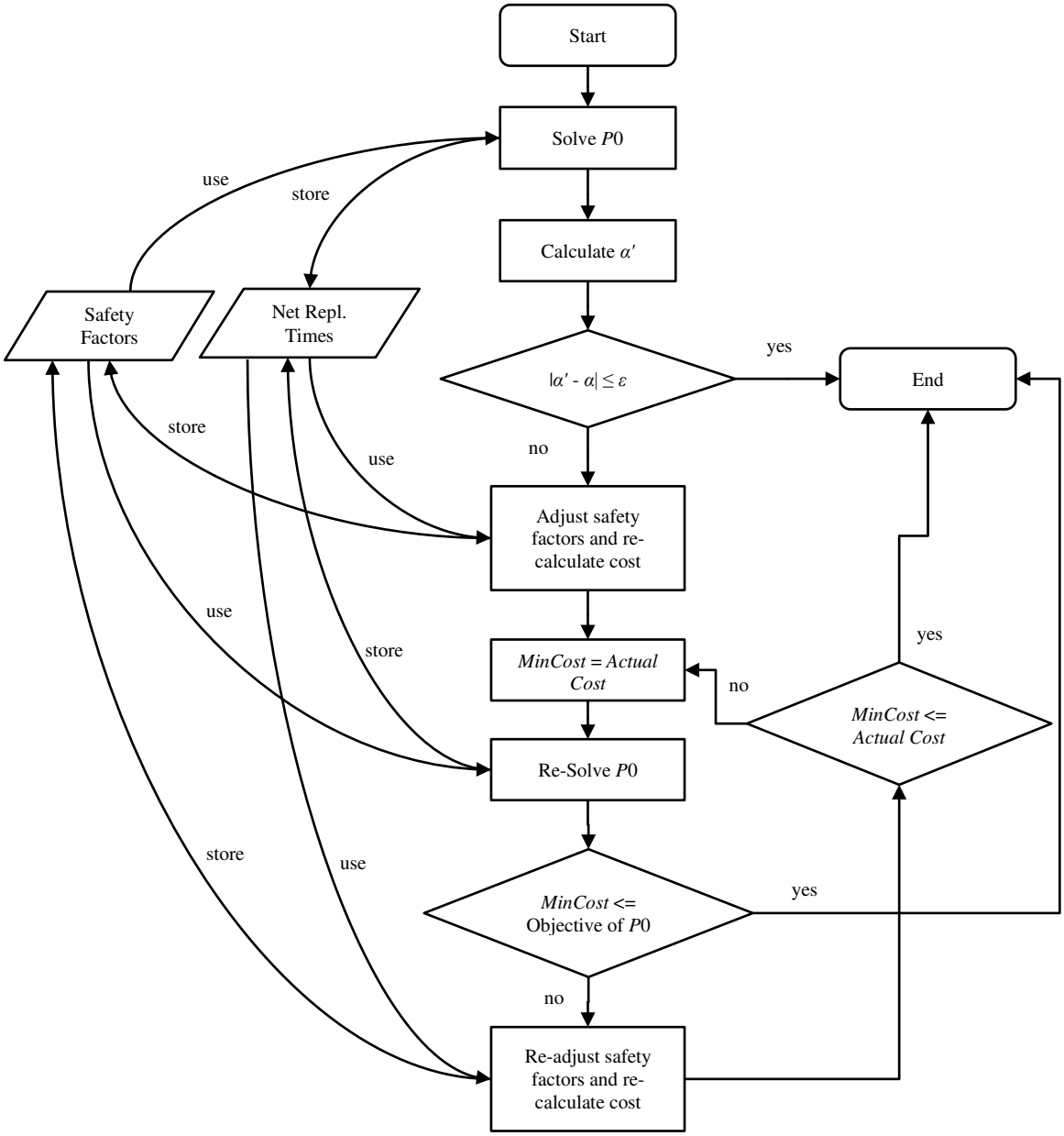

Figure 5.1: Flowchart diagram of Mitigation Approach II

5.1.3 Numerical Analysis

In this numerical analysis, we use the five-stage serial and real-world assembly test problems presented in Chapter 3 (Section 3.1.3). We compare the performance of Mitigation Approaches I and II in terms of computational time and solution quality evaluated based on the resulting safety stock cost increase.

Mitigation Approach I is implemented as follows. First, we model the problem *P*0 in GAMS 23.7 and use the solver BARON (version 9.3.1) to determine the net replenishment times and demand bounds associated with the initial solution of each test problem. Then we

assess the CSL deviation for this solution using a simulation model in Microsoft Excel. In order to determine the common safety factor that ensures the target CSL at the most downstream stage in our simulation model, we use the Goal Seek Tool in Microsoft Excel. This tool is based on a bisection method for which we set the tolerance level ε to 0.0001. It does not require specifying a maximum safety factor value $z^{\max}$ since it is already considered in the default settings of the tool.

In Mitigation Approach II, similarly to the first approach, we use the solver BARON to solve the problem *P*0 modelled in GAMS. We asses the CSL deviation for solutions obtained using our simulation model in Microsoft Excel. The safety stock adjustment procedure is coded in Microsost Excel VBA integrating the Goal Seek Tool. The tolerance level is set as $\varepsilon = 0.0001$ and the maximum safety factor $z^{\max} = 3.09$ corresponds to 99.9% CSL. For simplification, this numerical analysis is performed considering target service CSL values higher than 80%.

Five-Stage Serial System

In this subsection, the five-stage serial system with decreasing, uniform and increasing stage cost and lead time alternatives given in Section 3.1.3 of Chapter 3 is considered. Initial solutions are obtained solving the problem *P*0 under different safety factor values that relate to 80%, 85%, 90%, 95% and 99% target CSL. The CSL deviations are assessed using the simulation model under the demand bounds associated with these initial solutions. As presented in Section 3.1.4, the gap between the effectively observed and the target CSL are insensitive to demand parameters. There is no deviation if the initial solution implies that the safety stocks in the system are solely carried at the demand stage (i.e. at Stage 5). In this case, there is no need to use the mitigation approaches since the initial solution already ensures the target CSL at Stage 5. However, if a CSL deviation occurs for the initial solution, we mitigate this deviation applying the mitigation approaches proposed in Section 5.1.

The mitigation of the CSL deviation increases the total safety stock cost if the effectively observed CSL is less than the target one for the initial solution. Here, the solution

quality of the two mitigation approaches are evaluated in terms of the resulting safety stock cost increase. This is measured by $(Obj_{mit} - Obj_0)/Obj_0$ where Obj_0 is the total safety stock cost associated with the initial solution and Obj_{mit} is the total cost obtained as the outcome of the mitigation approach used. Table 5.1 and Table 5.2 summarise results obtained using Mitigation Approach I and Mitigation Approach II, respectively. Solutions that ensure the given target CSL at Stage 5, obtained with Mitigation Approach I, represent 6.3% safety stock cost increase on average. The cost performance of Mitigation Approach II is better than Mitigation Approach I since it leads to 2.6% cost increase on average. However, we cannot say that Mitigation Approach II dominates Mitigation Approach I in terms of cost. Counterexamples occur, e.g. for combinations of constant stage costs, decreasing lead times and 85%-99% target service levels.

For both approaches, the safety stock cost increase usually decreases when the target CSL increases. This may be related to the fact that the CSL deviation decreases for high target CSL values (see Section 3.1.4). The only exception is observed for the test problems with increasing stage costs and decreasing lead times. We note that this corresponds to the unique test problem where the net replenishment time of Stage 5 is the smallest among all stages (see Table 5.3 and Table 5.4).

Stage Cost	**Lead Time**	**Target Cycle-Service-Levels**				
		80.00%	*85.00%*	*90.00%*	*95.00%*	*99.00%*
Decreasing	Decreasing	0.00%	0.00%	0.00%	0.00%	0.00%
	Constant	0.00%	0.00%	0.00%	0.00%	0.00%
	Increasing	0.00%	0.00%	0.00%	0.00%	0.00%
Constant	Decreasing	10.67%	8.61%	7.26%	6.12%	4.28%
	Constant	16.58%	13.31%	10.37%	8.75%	5.20%
	Increasing	0.00%	0.00%	0.00%	0.00%	0.00%
Increasing	Decreasing	3.59%	3.90%	3.91%	4.75%	5.48%
	Uniform	15.15%	12.41%	10.33%	8.14%	5.08%
	Increasing	37.64%	30.35%	24.22%	19.26%	9.71%

Table 5.1: Safety stock cost increases that result from applying Mitigation Approach I

Stage Cost	Lead Time	Target Cycle-Service-Levels				
		80.00%	*85.00%*	*90.00%*	*95.00%*	*99.00%*
Decreasing	Decreasing	0.00%	0.00%	0.00%	0.00%	0.00%
	Uniform	0.00%	0.00%	0.00%	0.00%	0.00%
	Increasing	0.00%	0.00%	0.00%	0.00%	0.00%
Uniform	Decreasing	8.70%	8.70%	8.70%	8.70%	8.70%
	Uniform	1.64%	1.64%	1.64%	1.64%	1.64%
	Increasing	0.00%	0.00%	0.00%	0.00%	0.00%
Increasing	Decreasing	1.66%	2.06%	2.30%	2.76%	3.26%
	Uniform	15.74%	15.74%	4.58%	3.53%	2.69%
	Increasing	2.05%	2.05%	2.05%	2.01%	2.00%

Table 5.2: Safety stock cost increases that result from applying Mitigation Approach II

The net replenishment times (from Stage 1 to 5) associated with the initial solutions used for both approaches are listed in Table 5.2. We remind that a stage with strictly positive net replenishment time is a stage that holds safety stock. By the construction of Mitigation Approach I, the initial and final safety stock placement solutions are the same, i.e. safety stock placements in the system do not change after the mitigation of the CSL deviation. However, this is not the case for Mitigation Approach II since we evaluate several solutions through this mitigation procedure. Final solutions obtained with Mitigation Approach II are given in Table 5.4. This shows that in most of the cases, it is advantageous in terms of cost to hold safety stock solely at Stage 5 in order to prevent the CSL deviation. For this solution, the net replenishment time of Stage 5 equals its maximum replenishment time (100 periods) while the net replenishment time of other stages equals zero. The cost advantage of this solution may vanish when the per-unit holding cost is significant at Stage 5, i.e. in case of increasing stage costs.

In terms of computational time, Mitigation Approach I is better than Mitigation Approach II for all test problems. The relevant solutions are obtained within 30 seconds with the former one while the later one requires 4 minutes on average.

Stage Cost	Lead Time	Initial Solution
Decreasing	Decreasing	(0,0,0,0,100)
	Uniform	(0,0,0,0,100)
	Increasing	(0,0,0,0,100)
Uniform	Decreasing	(36,0,0,0,64)
	Uniform	(20,0,0,0,80)
	Increasing	(0,0,0,0,100)
Increasing	Decreasing	(36,28,20,0,16)
	Uniform	(20,20,0,0,60)
	Increasing	(4,12,0,0,84)

Table 5.3: Net replenishment times of initial solutions used for Mitigation Approaches I and II

Stage Cost	Lead Time	Final Solution				
		80.00%	*85.00%*	*90.00%*	*95.00%*	*99.00%*
Decreasing	Decreasing	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)
	Uniform	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)
	Increasing	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)
Uniform	Decreasing	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)
	Uniform	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)
	Increasing	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)
Increasing	Decreasing	(36,28,0,32,4)	(36,28,0,32,4)	(36,28,0,32,4)	(36,28,0,32,4)	(36,28,0,32,4)
	Uniform	(0,0,0,0,100)	(0,0,0,0,100)	(20,0,40,0,40)	(20,0,40,0,40)	(20,0,40,0,40)
	Increasing	(0,0,0,0,100)	(0,0,0,0,100)	(0,0,0,0,100)	(4,0,0,0,96)	(4,0,0,0,96)

Table 5.4: Net replenishment times of final solutions obtained by Mitigation Approach II

Real-World Assembly System

In this subsection, we consider the real-world assembly system presented in Section 3.1.3 of Chapter 3. The CSL deviations under different target CSL values are given in Section 3.1.4. Here, we mitigate the CSL deviations using Mitigation Approaches I and II. The total safety stock costs obtained for different target CSL values ranging from 80% to 99% are displayed in Figure 5.2. The upper, middle and lower lines corresponds to solutions obtained by Mitigation Approach I, Mitigation Approach II and the original GSM (initial solution obtained by solving the problem *P*0), respectively.

We notice that for this real-world example, Mitigation Approach II outperforms Mitigation Approach I in terms of cost. The average cost increases resulting from Mitigation Approaches I and II are 37.3% and 8.9%, respectively. However, Mitigation Approach I is

quiet better in terms of computational time than Mitigation Approach II. The time required to obtained the relevant solutions with Mitigation Approach I is less than 30 seconds while the computational time of Mitigation Approach II is 25 minutes on average. Compared to results obtained for the serial test problems, Mitigation Approach II performs better than Mitigation Approach I in terms of cost. However, the computational time of Mitigation Approach II increases significantly for assembly test problems, i.e. when the complexity of the supply chain network increases.

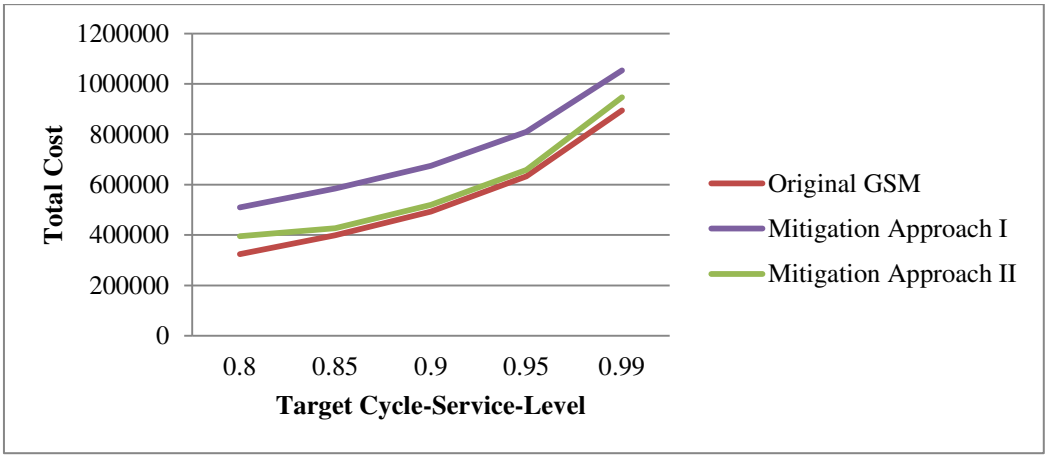

Figure 5.2: Total safety stock costs as a function of target CSL for the bulldozer supply chain

Let illustrate solutions obtained for the test problem with the target CSL of 95% which corresponds to the example considered by Graves and Willems (2003). Table 5.5 and Table 5.6 give the strictly positive net replenishment times and the safety factors obtained by using Mitigation Approaches I and II, respectively. The safety factor applied in the initial solution of this test problem is $z = \Phi^{-1}(0.95) = 1.645$. The total safety stock cost associated with this solution is 632,719$. As presented in Section 3.1.4 of Chapter 3 this solution represents a CSL deviation of 7.71%. The solution obtained by Mitigation Approach I mitigates the CSL deviation by applying a common safety factor $\hat{z} = 2.10$. This leads to an increase of 27.9% in the total safety stock cost. The solution obtained by Mitigation Approach II adjusts safety factors according to the net replenishment times and per-unit holding costs of stages. This leads to different safety factors for different stages. Besides, the safety stock placements (and hence the net replenishment times) associated with the final

solution is different from the initial solution. This solution represents a safety stock cost increase of 4% in comparison with the initial solution. This detailed illustration could be presented for any other value of target CSL.

Stage	Optimal Net Repl. Times (days)	Safety Factors
Case	15	2.10
Case & Frame	1	2.10
Fans	2	2.10
Final Assembly	32	2.10
Frame Assembly	19	2.10
Pin Assembly	14	2.10

Table 5.5: The final solution obtained by Mitigation Approach I

Stage	Optimal Net Repl. Times (days)	Safety Factors
Case	15	3.09
Fans	1	3.09
Final Assembly	33	1.67
Frame Assembly	19	2.43
Pin Assembly	13	2.81

Table 5.6: The final solution obtained by Mitigation Approach II

As said earlier, Minner (2000) mitigates the CSL deviation by adjusting the safety stock levels for a three-stage serial system. His approach consists of sequentially increasing safety stock levels by focusing on two-stage service level interactions and extreme strategies such as increasing only the upstream or only the downstream safety stock level. For his illustrative example increasing only the downstream safety stock is better in terms of cost compared to increasing only the upstream safety stock. This is because; the safety stock increase at the upstream stage cannot be compensated by holding cost advantages for this illustrative example. For the test problems that we present in this section, we obtain opposite results. The per-unit holding cost at the most downstream stage (Final Assembly) is significant compared to those of upstream stages (see Table 3.4, Section 3.1.3). Hence,

increasing safety factors at the upstream stages where the per-unit holding costs are quiet low gives significantly better results. This shows that the cost performance of different startegies proposed depends greatly on the input data. However, it is not obvious how to fairly compare the approaches that we propose with those of Minner (2000) since he is based on a different setting.

5.2 Cycle-Service-Level Deviation in Decentralised Supply Chains

In this section, we present some issues related to the CSL deviation in decentralised systems. In a decentralised system, different autonomous actors control different parts of the supply chain. These actors may have competing and conflicting interests and objectives. This raises some specific problems of coordination. The literature on the supply chain coordination problem in decentralised systems is large and growing (see, e.g., Cachon, 2003; Jemai and Karaesmen, 2007; Hennet and Arda, 2008; Hennet, 2009). Recently, researchers showed that the GSM is well suited for decentralised systems. For decentralised systems that operate according to the GSM setting, researchers are interested in coordinating the service time to be applied between different actors and the price paid by a downstream actor for each unit ordered to an upstream actor. For instance, Schoenmeyr (2008) proposes a simple contract structure in order to facilitate the relationship between two actors. The proposed contract may realign the actors' incentive structures and enable the supply chain to operate according to the globally optimal solution. Similarly, Egri (2012) studies the GSM approach in a decentralised supply chain involving a number of autonomous stages and presents different mechanisms that facilitate the alignment of conflicting goals in order to achieve the globally optimal solution. To the best of our knowledge, how to specify the demand bounds at different parts of the supply chain is not discussed in the existing literature. Existing works ignore the conflict that may occur due to the incompatible demand bounds used by different actors of the supply chain. However, the issue of customer service level deviation presented in Section 3.1 of Chapter 3 becomes even more important in decentralised supply chains. This is because, in a decentralised supply chain, the downstream actors would probably not be informed of the demand bounds applied by the upstream actors. Thus, they would not be

able to predict and mitigate the resulting CSL deviation at their most downstream stage. Indeed, the safety factors (and hence the demand bounds) to be applied by different actors should be included into the bargaining process in order to deal with this issue. This section presents a first step in understanding the issue related to customer service level deviation in decentralised supply chains.

We mainly base on the results developed by Schoenmeyr (2008) who shows that provided that the actors can agree on the right service time between them, it will be in their best interest to operate the supply chain according to the globally optimal safety stock placement solution. However, he considers safety factors as exogenously specified parameters that reflect the external customer service level of the downstream actor. This neglects the CSL deviation that would occur in such systems. Indeed, the CSL deviation observed at the downstream actor can be mitigated by using the mitigation approaches proposed in Section 5.1. However, if both actors apply a mitigation approach independently, the downstream actor may still face the CSL deviation issue. In this setting, we aim at investigating the CSL deviation faced by the most downstream actor. Section 5.2.1 discusses assumptions used for this analysis while Section 5.2.2 summarises our numerical study.

5.2.1 Assumptions

We focus on systems that are controlled by two different actors. We assume that both actors apply the GSM to their own parts of the supply chain and mitigate the CSL deviation for their own parts. As in Schoenmeyr (2008), we assume that both actors agree with the service time which corresponds to the optimal solution that is obtained by considering safety factors as exogenously specified parameters and by solving the problem $P0$ for the whole supply chain. Based on this initial solution, we assume that both actors mitigate the CSL deviation independently using a mitigation approach. That is, both actors guarantee to ensure the CSL required by the external customer of the downstream actor considering their own demand bounds. We note that the use of Mitigation Approach I by both actors does not affect the service time quoted between them since the safety stock placement solution will not change after this mitigation. This makes the assumption of Schoenmeyr (2008) still valid.

However, if at least one actor uses Mitigation Approach II and observes that the best safety stock placement solution is different after mitigation, they might finally agree on a different service time than the service time proposed by Schoenmeyr (2008). Since we base on the results presented by Schoenmeyr (2008), we only consider the case where both actors use Mitigation Approach I.

5.2.2 Numerical Analysis

The purpose of this numerical analysis is to measure the CSL deviation observed at the downstream actor when both actors mitigate the CSL deviation for their own parts in the supply chain.

We use the five-stage serial and the five-echelon assembly test problems presented in Chapter 3 (Section 3.1.3). We consider that these systems are controlled by two actors. We investigate cases when the upstream actor (Actor 1) controls the most upstream 1, 2, 3, or 4 echelons. We consider that the external customer of the downstream actor (Actor 2) asks for 80%, 85%, 90%, 95% and 99% service levels which define the target CSL of both actors.

As in Schoenmeyr (2008) we assume that both actors agree with a safety stock placement solution that is equivalent to the optimal solution obtained by solving the problem *P*0. Hence, given this safety stock placement solution and the relevant target service levels, we apply Mitigation Approach I within the parts of both actors as presented in Section 5.1.3. We then asses the CSL deviation at the most downstream stage of Actor 2 under the resulting safety factors and demand bounds.

Five-Stage Serial System

This subsection presents the results obtained for the five-stage serial test problems. After independently applying Mitigation Approach I for both actors, the effectively observed CSL at the demand stage (Stage 5) of Actor 2 is denoted by α_{mit}. The effectively observed CSL α_{mit} is calculated using (3.3) under the demand bounds applied in the whole system. Table 5.7 summarises the relative gaps $(\alpha - \alpha_{mit})/\alpha$ that are averaged for different target CSL values.

For these serial test problems, the CSL deviation at Stage 5 increases when the number of stages controlled and the number of demand bounds applied by the upstream actor increase. In accordance with the solutions obtained in Section 3.1.4, the worst case is observed for increasing stage cost and increasing lead time combinations. The CSL deviation observed in these serial test problems is 1.4% on average.

Stage cost	**Lead time**	**Initial/Final Solution**	**Upstream Actors' Most Downstream Stage**			
			1	*2*	*3*	*4*
Decreasing	Decreasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%
	Uniform	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%
	Increasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%
Uniform	Decreasing	(36,0,0,0,64)	1.84%	1.84%	1.84%	1.84%
	Uniform	(20,0,0,0,80)	2.82%	2.82%	2.82%	2.82%
	Increasing	(0,0,0,0,100)	0.00%	0.00%	0.00%	0.00%
Increasing	Decreasing	(36,28,20,0,16)	0.16%	0.40%	0.77%	0.77%
	Uniform	(20,20,0,0,60)	2.70%	2.70%	2.70%	2.70%
	Increasing	(4,12,0,0,84)	4.07%	5.14%	5.14%	5.14%

Table 5.7: CSL deviation results for the decentralised serial system

Real-World Assembly System

We perform the same analysis for the real-world assembly system presented in Section 3.1.3. Table 5.8 summarises the relative gaps for different target CSL values. Similarly to what is observed in the serial test problems, the CSL deviation usually increases when the number of stages controlled and the number of demand bounds applied by the upstream actor increase (the case of 99% CSL target represents the only counterexample). For these assembly test problems, the relative gap between the effectively observed and the target CSL is 8.1% on average.

This analysis shows that the CSL deviation may represent an important issue in real-world decentralised supply chains that operate according to the GSM setting. If each actor mitigates the CSL deviation by solely considering their own demand bounds, most downstream stage can still face a deviation in its service level. This is due to the non-consideration of the dependencies between the demand bounds applied at the upstream actor

and the effectively observed CSL at the most downstream stage that faces final customer demand. However, these results present the case where the mitigation of the CSL deviation does not affect the initially agreed service time between the actors. Indeed, the safety factors to be applied should be considered while bargaining over the service time to be quoted between different actors.

Target CSL	**Upstream Actors' Most Downstream Echelon**			
	1	*2*	*3*	*4*
99%	1.25%	0.93%	0.93%	0.93%
95%	4.23%	4.24%	4.24%	4.24%
90%	6.62%	8.58%	8.58%	8.58%
85%	8.31%	12.91%	12.91%	12.91%
80%	9.66%	17.19%	17.19%	17.19%

Table 5.8: CSL deviation results for the decentralised assembly system

5.3 Conclusion

In the first part of this chapter, we showed how to mitigate the CSL deviation in the GSM setting by adjusting the safety factors applied at different stages of the supply chain. We proposed two mitigation approaches and compared their performances in terms of total cost and computational time. For a real-world assembly system considered for the numerical analysis, the first approach is better than the second one in terms of computational time (30 seconds vs. 25 minutes) while the second approach provides better solutions in terms of cost (8.9% vs. 37.3% cost increase). The mitigation of the CSL deviation increases the total safety stock cost and the magnitude of this increase depends greatly on the net replenishment time of stages. Hence, a better solution in terms of cost can represent a different safety stock placement solution than the one initially obtained by solving the problem $P0$. Further research can be conducted in order to take into account the cost impact of the mitigation through the solution procedure of the GSM. In other words, the relevant safety stock adjustments can be integrated into the solution procedures proposed for the original GSM (e.g., Minner, 2000; Graves and Willems, 2000; Humair and Willems, 2011). This would

enable to obtain the optimal safety stock placement solution that minimises the total safety stock cost while achieving the target service level at the most downstream stage.

In the second part of this chapter, we focused on decentralised systems where different parts of the supply chain are controlled by different actors. In such systems, even the CSL deviation is mitigated for each actor, the most downstream actor may still face a deviation in its service level due to the demand bounds applied at its upstream actors. We measure the relevant deviation considering five-echelon serial and real-world assembly systems controlled by two actors. Results presented for the real-world system show that the resulting relative deviation is 8.1% on average. This shows that including the safety factors into the bargaining process is important since the CSL deviation may be significant when different actors mitigate the CSL deviation independently. However, under the proposed setting, different actors might not be able to agree with the service time which is considered as the best by Schoenmeyr (2008). Our work represents the first step in understanding the complications that may arise in decentralised supply chains. A further step is to propose a contract structure that facilitates the relationship between different actors that bargain over the safety factors to be applied at different parts of the supply chain.

GENERAL CONCLUSION AND FUTURE RESEARCH DIRECTIONS

Multi-echelon supply chains consist of several stages associated with procurement, manufacturing and transportation processes. For supply chains facing external demand uncertainty, the development of computationally tractable approaches for dealing with the multi-echelon inventory optimisation problem is a complex task. In this thesis, we developed several contributions to this problem using the Guaranteed-Service Model (GSM) approach due to the economic, computational and practical advantages of this approach. First of all, we conducted a comprehensive literature review which gives a synthesis of the various works developed so far. Hence, we identified the gaps in the GSM literature that are being addressed in this thesis. For instance, we showed that some specific assumptions of the GSM may have a significant impact on customer service levels and safety stock costs. In addition, we developed an extension of the GSM that enables to simultaneously optimise the reorder intervals and order-up-to levels in general acyclic multi-echelon systems and we proposed an optimisation procedure that enables to obtain near optimal solutions with reasonable computational time for this model. Finally, we studied the issue of Cycle-Service-Level (CSL) deviation under the GSM assumptions and proposed two approaches in order to mitigate this deviation. Concerning these mitigation approaches, the second approach is better than the first in terms of cost performance while the first one outperforms the second one in terms of computational time. We also presented some issues related to the CSL deviation in decentralised supply chains.

In addition to the concluding sections of the previous chapters, several interesting directions remain for future consideration.

This thesis considers the common GSM assumption that indicates that when external demand exceeds the pre-specified demand bounds; the excess demand is handled outside the normal supply chain, by using some extraordinary measures such as overtime production, express expediting and subcontracting. However, we do not explicitly model the impact of extraordinary measures on the customer service level and total cost. In the literature, only few

contributions are provided concerning the integration of extraordinary measures into the GSM. Existing models either lost the computational advantages of the GSM that stems from its deterministic structure (Rambau and Schade, 2010) or focus on simple supply chain structures such as serial, assembly or two-echelon distribution systems introducing express expediting as extraordinary measure to use (Klosterhalfen and Minner, 2010; Li, 2013). Further research can be pursued in order to integrate the impact of different extraordinary measures into the models that we present in this thesis. Besides, similarly to almost all models in the GSM literature, we use the CSL service measure while specifying the demand bounds. Extensions of these results to other types of service measures commonly used in practice (e.g. fill-rate service measures) seems worthwhile.

In this thesis, external customer demand is assumed to be stationary. In practice, many companies experience non-stationary demand due to short product life cycles, seasonality, sales-force incentives etc. The extensions of the GSM to non-stationary demand are developed in the literature (see, e.g., Graves and Willems, 2008; Neale and Willems, 2009). These results can be used to consider non-stationary demand for the models studied in this thesis. However, as demonstrated by Graves and Willems (2008), constant service times may result in sub-optimal solutions for this case. In the literature, a general idea on the cost performance of constant service time solutions has not been provided yet. The consideration of dynamic service times can also be worthwhile in order to determine better solutions for systems facing non-stationary demand and to quantify the sub-optimality of constant service time solutions. However, from a practical point of view, a dynamic service time model may cause several issues since the safety stock locations will change dynamically rendering the model and its implementation more complicated.

This thesis solely focuses on periodic-review, installation order-up-to policies. Taking into consideration the specific environment in which the studied systems operate, other inventory control policies could be more appropriate for the considered systems. Besides, it is shown in the literature that the standard order-up-to policies are not necessarily optimal in terms of cost for guaranteed-service supply chains (Schoenmeyr, 2008). Although the determination of an optimal inventory control policy is challenging for general multi-echelon

systems, other inventory control policies that perform better than periodic-review order-up-to policies (in terms of cost) can be developed following for instance the ideas proposed by Schoenmeyr (2008) and Shi and Zhao (2010).

We limit our attention to serial and assembly systems in Chapter 3 and Chapter 5. The extension of results provided in these chapters to distribution and general acyclic structures requires the consideration of the stock allocation problem encountered in multi-echelon systems operating with periodic-review policies (see, e.g., Jackson, 1988; Graves, 1996; Marklund and Rosling, 2012). One can incorporate existing approaches that enable to deal with the relevant stock allocation problem in order to extend our results. The problem presented in Chapter 4 is modelled for general acyclic systems. Further research can be conducted to extend the model proposed to even more complex systems. In the literature solution techniques for general cyclic multi-echelon systems have not been provided yet despite the requirements in real-world applications, e.g., in chemical and pharmaceutical industries. The consideration of cycles in the supply chain system reveals additional complexity and represents another challenging future research direction.

RESUME ETENDU

Une chaîne logistique est un système qui intègre plusieurs processus pour transformer les matières premières approvisionnées auprès de fournisseurs externes en un produit final à livrer à des clients externes. De nombreuses chaînes logistiques peuvent être caractérisées comme de larges et complexes systèmes multi-échelons, pouvant être constitués de plusieurs milliers d'étages (cf. Willems, 2008). Dans ces systèmes multi-échelons, chaque étage est associé à un processus tel que l'approvisionnement d'une matière première, la fabrication d'un composant, l'assemblage d'un produit final, son transport à partir d'un centre de distribution central vers un entrepôt régional ou d'un entrepôt régional vers un magasin (Graves et Willems, 2000). L'un des enjeux majeurs associé au management de ces systèmes est la gestion efficace des stocks lorsque la demande est incertaine, les coûts de stockage sont importants et les exigences en terme de niveau de service client sont élevées. Cela nécessite en particulier de spécifier les niveaux de stocks aux différents étages afin de minimiser le coût total du système et de satisfaire les niveaux cibles de service client.

Dans cette thèse, nous nous concentrons principalement sur les problématiques liées à l'optimisation de stocks de sécurité dans les systèmes multi-échelons. Le stock de sécurité est considéré comme un levier pour couvrir les incertitudes dans les systèmes de stocks. Dans un système de stocks multi-échelons, le niveau de stock de sécurité à chaque étage doit être optimisé de manière appropriée afin d'atteindre les niveaux cibles de service client à moindre coût. En effet, le niveau de stock de sécurité local associés à chaque étage peut être déterminée de façon indépendante en utilisant les modèles de stocks mono-échelon qui ont été largement étudiés à ce jour (cf. par ex., Silver et al., 1998; Zipkin, 2000). Cependant, une telle approche ne considère que les paramètres associés à l'étage correspondant. Elle entrainerait donc des stocks de sécurité redondants en raison de la non-prise en compte des interdépendances concernant les coûts et les niveaux de service des étages reliés. L'approche de l'optimisation de stocks de sécurité multi-échelons vise à optimiser les stocks de sécurité avec une vision globale, en considérant simultanément tous les étages de la chaîne logistique, du fournisseur externe au client externe. Bien que l'approche multi-échelons impose des défis

de calcul importants, elle donne de meilleurs résultats par rapport à l'approche mono-échelon en termes de coût et de niveau de service client. La Figure I illustre les approches mono et multi-échelons[2].

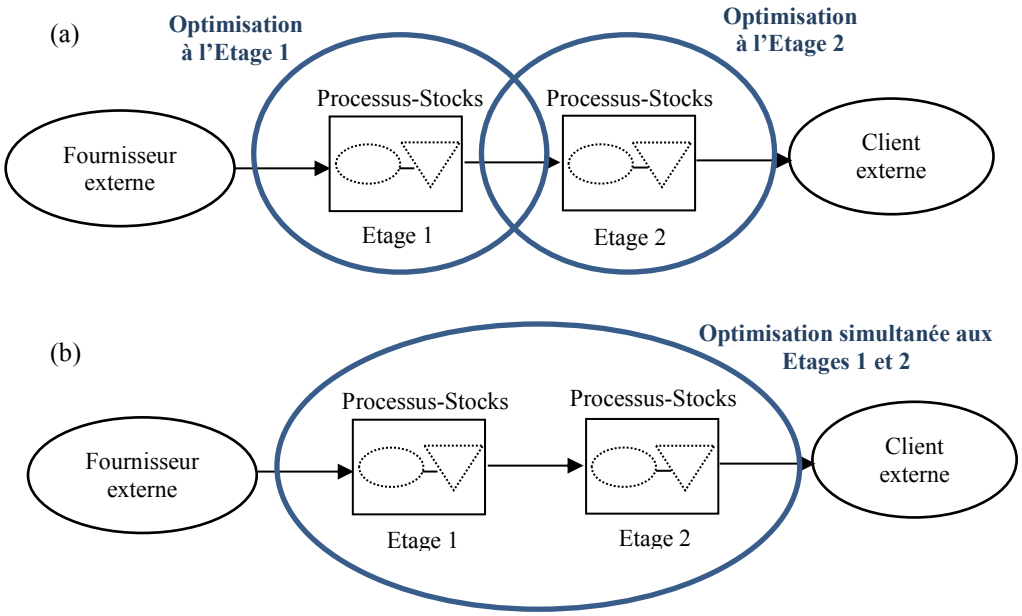

Figure I: (a) Approche mono-échelon, (b) Approche multi-échelons

Selon Simchi-Levi et Zhao (2012), trois raisons ont contribué à l'essor de l'approche d'optimisation de stocks multi-échelons: (1) la disponibilité des données relatives à la demande et aux délais de réapprovisionnement, (2) les motivations de l'industrie à utiliser les méthodes scientifiques pour la gestion de stocks, (3) les développements récents en modélisation et en conception d'algorithmes pour traiter les structures multi-échelons générales. Le rapport du benchmark du Groupe Aberdeen (2007) confirme également la motivation des entreprises à utiliser les outils d'optimisation de stocks multi-échelons. Selon ce rapport, en 2007, les 210 entreprises interrogées placent l'optimisation des stocks au premier rang comme domaine d'investissement où l'optimisation des stocks multi-échelons représente leur priorité absolue. La mise en œuvre des outils d'optimisation de stocks multi-échelons peut permettre une amélioration du niveau de service de 3,1% et une diminution du cycle de trésorerie de 15% (Aberdeen Group, 2012).

[2] Cette figure est inspirée de Klosterhalfen (2010)

L'approche d'optimisation de stocks multi-échelons a été largement étudiée dans la littérature. Les chercheurs ont proposé le Modèle de Service Stochastique (SSM) et le Modèle de Service Garanti (GSM) qui sont introduits, respectivement, par Clark et Scarf (1960) et Simpson (1958). Ces deux approches diffèrent en termes d'hypothèses utilisées concernant la gestion des variations de la demande et les temps de service. L'approche du SSM se concentre principalement sur les systèmes du type série, assemblage ou distribution à deux-échelons. Par conséquent, le déploiement du SSM dans l'industrie est relativement limité. Cependant, les hypothèses simplificatrices faites dans l'approche du GSM, permettent d'étudier les chaînes logistiques de taille réelle.

L'approche du GSM a suscité un grand intérêt au cours de ces dix dernières années. Les extensions académiques du GSM ont rendu cette approche plus réaliste. En parallèle, les méthodes de résolution efficaces ont permis de faire face aux systèmes multi-échelon actuels de grande taille ayant une structure complexe. Ainsi, l'approche du GSM a permis de réaliser des bénéfices importants dans l'industrie. Par exemple, Billington et al. (2004) ont montré que les économies réalisées en utilisant l'approche du GSM pour les produits de Hewlett-Packard tels que les appareils photos numériques et les consommables jets d'encre ont dépassé 130 M$. Farasyn et al. (2011) rapportent que les modèles multi-échelons basés sur l'approche du GSM produisent 7% de réduction des stocks en moyenne dans les unités commerciales de Procter & Gamble. Wieland et al. (2012) décrivent un projet d'optimisation de stocks multi-échelons chez Intel et indiquent que, après sa mise en œuvre, les niveaux de stocks ont été réduits de plus de 11% tout en fournissant des niveaux de service supérieurs à 90% en moyenne. Les avantages pratiques, computationnels et économiques de l'approche du GSM sont reconnus à la fois par les scientifiques et par les professionnels.

Dans cette thèse, nous traitons le problème d'optimisation de stocks de sécurité multi-échelons en se concentrant particulièrement sur l'approche du GSM. Nous fournissons plusieurs contributions dans ce sens. Cette thèse est divisée en 5 chapitres. Après une introduction générale, le **Chapitre 1** présente les fondamentaux qui forment la base des chapitres à venir. Il comprend la terminologie de base et la présentation détaillée des deux

modèles principaux, le GSM et le SSM proposés afin de traiter le problème d'optimisation de stocks de sécurité multi-échelons.

Notre première contribution dans cette thèse est de fournir une revue de la littérature exhaustive sur l'approche du GSM. Cette revue est présentée dans le **Chapitre 2**. En effet, les littératures appartenant à l'approche du SSM et du GSM ne sont pas au même degré de maturité. La littérature sur l'approche du GSM est vaste et contient de nombreux travaux réalisés depuis celui de Clark et Scarf (1960). D'excellentes revues de littérature sont également présentées dans cette littérature (cf. par ex. Diks et al., 1996; Axsäter, 2003; Simchi-Levi et Zhao, 2012). D'autre part, en comparaison avec l'approche du SSM qui est largement étudiée dans la littérature, la recherche sur l'approche du GSM a suscité un intérêt particulier dans la dernière décennie. Au meilleur de notre connaissance, nous n'avons pas relevé de revue de littérature proposant une synthèse des différents travaux réalisés à ce jour. Nous présentons une revue exhaustive en classant ces travaux en trois axes: les hypothèses de modélisation considérées, les méthodes de résolution développées et les applications industrielles/résultats obtenus. Notre revue de littérature nous permet d'identifier certaines lacunes dans la littérature du GSM et nous mène à des questions de recherche considérées dans cette thèse. Une version préliminaire de ce travail est publiée dans les actes de la conférence internationale «*14th IFAC Symposium on Information Control Problems in Manufacturing, INCOM'12 »* (Eruguz et al., 2012).

De notre revue de littérature, nous constatons que l'impact de certaines hypothèses spécifiques du GSM sur le coût total et sur le niveau de service client n'est pas étudié de manière détaillée. En particulier, une analyse approfondie est nécessaire pour comprendre l'impact des hypothèses telles que la demande bornée, les temps de service garanti et les périodes d'approvisionnement en communes. Par conséquent, notre deuxième contribution dans cette thèse est de fournir une telle analyse en étudiant l'impact de chaque hypothèse séparément. Cette analyse est présentée dans le **Chapitre 3**. Dans cette étude, nous nous concentrons principalement sur les systèmes du type série et assemblage. Pour l'analyse numérique de chaque hypothèse, nous considérons les mêmes problèmes de test associés à un

système de cinq-échelons en série présentés précédemment par Schoenmeyr (2008) et un système d'assemblage réel examiné par Graves et Willems (2003).

La première hypothèse examinée dans le Chapitre 3 indique que la demande est bornée à chaque étage de la chaîne logistique. En pratique, la borne de demande dans un étage représente la quantité maximale de demande qui peut être satisfaite à partir des stocks de cet étage durant un certain temps de couverture. Dans la littérature, les bornes de demande sont généralement spécifiées en utilisant un facteur de sécurité qui se rapporte à un Niveau de Service-par-Cycle (CSL) ciblé. Notre analyse montre que le CSL effectivement observé dans un étage qui fait face à la demande du client externe serait habituellement inférieur à celui ciblé. Sous différents CSL ciblés pour le système du monde réel qu'on considère, l'écart relatif est de 25% en moyenne.

La deuxième hypothèse examinée dans le Chapitre 3 implique que chaque étage promet un temps de service garanti à ses clients et offre un niveau de service de 100% pour le temps de service proposé. Par conséquent, les commandes en attente ne sont pas autorisées entre les étages amont-aval. En effet, le coût total de stocks de sécurité obtenus sous cette hypothèse peut être considérablement supérieur à une solution que l'on peut obtenir sans cette hypothèse. Pour le système du monde réel considéré dans notre analyse numérique, l'hypothèse de temps de service garanti entraîne une augmentation de coûts de stocks de sécurité de 42,2 %.

La troisième hypothèse considérée dans le Chapitre 3 concerne la politique d'approvisionnement du GSM. Le GSM de base suppose que chaque étage fonctionne avec une politique à recomplétement périodique avec une période d'approvisionnement commune pour tous les étages. En pratique, les périodes d'approvisionnement peuvent différer d'un étage à l'autre du fait des économies d'échelle et/ou des disponibilités des ressources. Nous montrons comment intégrer dans le GSM les périodes d'approvisionnement dépendantes et imbriquées (« nested » en anglais). Cela nécessite le développement d'une expression appropriée pour les bornes de demande. Notre analyse numérique montre que le GSM de base peut entrainer une augmentation significative du coût total de stocks quand les périodes d'approvisionnement sont longues. Sous différents profils de période d'approvisionnement

considérés dans l'analyse numérique, l'augmentation du coût total de stocks de sécurité a pu atteindre 19,1%.

La première partie de l'analyse réalisée afin de quantifier l'impact des hypothèses du GSM (hypothèse de la demande bornée) est publiée dans les actes de la conférence internationale «5th International Conference on Modeling, Simulation and Applied Optimization, ICMSAO'13 » (Eruguz et al., 2013c). Une version préliminaire de cette étude est présentée dans l'école d'été « 11th ISIR Summer School on Research Trends in Inventory Management and Modeling » (Eruguz et al., 2013d). Les conséquences liées aux hypothèses étudiées démontrent qu'il existe un besoin pour de nouveaux développements des modèles concernant l'approche du GSM.

Dans la littérature du GSM, les modèles existants considèrent les périodes d'approvisionnement des étages en tant que données d'entrée du problème d'optimisation de stocks de sécurité multi-échelons. Notre troisième contribution dans cette thèse est de fournir une extension du GSM en intégrant les coûts d'approvisionnement fixes dans le modèle afin d'optimiser les niveaux de récomplétement et les périodes d'approvisionnement simultanément. Nous présentons cette contribution dans le **Chapitre 4**. Dans cette étude, nous nous concentrons sur les périodes d'approvisionnement imbriquées et en Puissance de Deux (PO2) en raison des avantages pratiques et computationnels des politiques associées (cf. par ex., Maxwell and Muckstadt, 1985; Muckstadt et Roundy, 1993; Yao et Wang, 2006). Sous une politique de PO2 imbriquée, les périodes de réapprovisionnements sont des puissances de deux et la période d'approvisionnement d'un étage ne peut pas être plus grande que celles de ses étages en amont. Afin de calculer les bornes de demande sous ces conditions, nous nous sommes inspirés de l'expression que nous avons proposée lors de l'analyse de l'hypothèse des périodes d'approvisionnement communes dans le Chapitre 3. Sous l'existence de ces bornes de demande, nous proposons d'abord un modèle déterministe de Programmation Non-Linaire en Nombres Entiers (NLIP) qui détermine à la fois les niveaux de recomplétement et les périodes d'approvisionnement PO2 dans les systèmes multi-échelons acycliques généraux. Deuxièmement, en définissant des bornes raisonnables pour les variables de décision du modèle de NLIP, nous proposons une approche directe

améliorée qui réduit le temps de calcul tout en résolvant le modèle de NLIP pour obtenir des solutions optimales globales. Troisièmement, nous proposons une Procédure d'Optimisation Séquentielle (SOP) pour obtenir des solutions proches de l'optimal avec un temps de calcul raisonnable. L'analyse numérique montre que pour un système multi-échelons acyclique général avec des paramètres générés aléatoirement, le SOP est en mesure d'obtenir des solutions proches de l'optimal avec un écart d'optimalité de 0,46 % en moyenne en quelques secondes. Ce travail a été publié dans la revue internationale « International Journal of Production Research » (Eruguz et al., 2014) .

Notre dernière contribution qui porte sur l'atténuation de l'écart de CSL dans le cadre du GSM est présentée dans le **Chapitre 5**. L'analyse sur l'hypothèse de la demande bornée dans le Chapitre 3 montre que le CSL effectivement observé par le client final peut être inférieur à celui ayant été ciblé. Dans la première partie de cette étude, nous montrons comment réduire l'écart de CSL en ajustant les facteurs de sécurité appliqués à différents étages de la chaîne logistique. Nous proposons deux approches d'atténuation et comparons leurs performances en termes de coût total et temps de calcul. L'étude numérique réalisée sur un système du monde réel montre que la première approche est plus performante que la seconde en termes de temps de calcul (30 secondes contre 25 minutes) tandis que la seconde approche offre des meilleures solutions en termes de coût (augmentation des coûts de 8,9% contre 37,3%). Dans la deuxième partie de cette étude, nous nous concentrons sur les systèmes décentralisés où les différentes parties de la chaîne logistique sont contrôlées par des acteurs différents. Dans de tels systèmes, même si l'écart de CSL est atténué par chaque acteur, les acteurs les plus en aval peuvent encore faire face à un écart de CSL en raison des bornes de demande appliquées chez l'acteur en amont. Nous montrons que l'écart de CSL peut être important (8,1% d'écart en moyenne pour le système d'assemblage présenté par Willems (2003) et contrôlé par deux acteurs). Une version préliminaire de cette étude est acceptée pour publication dans les actes de la conférence internationale « 5th International Conference on Industrial Engineering and Systems Management, IESM'13 » (Eruguz et al., 2013a) et présentée à la conférence internationale « MSOM INFORMS Conference 2013 » (Eruguz et al., 2013b).

En conclusion, cette thèse contribue au problème d'optimisation de stocks multi-échelons en se basant sur l'approche du GSM. Plusieurs pistes intéressantes restent à explorer ultérieurement telles que l'intégration de l'impact des mesures extraordinaires à utiliser en cas de demande en excès, la considération de la demande non-stationnaire, des nouvelles politiques de gestion de stocks (suivant par ex. les pistes données par Schoenmeyr, 2008 et Shi et Zhao, 2010) et des structures encore plus complexes comme les systèmes généraux cycliques. Cette thèse ouvre de futures pistes de recherche à la fois intéressantes et difficiles méritant d'être explorer.

References

Aberdeen Group, 2012. *Inventory optimization – impact of a multi-echelon approach*, Boston: Aberdeen Group.

Aberdeen Group, 2007. *The supply chain innovator's technology footprint 2007 - a benchmark report on companies' technology investment plans for gaining immediate and strategic payback*, Boston: Aberdeen Group.

Axsäter, S., 1997. On deficiencies of common ordering policies for multi-level inventory control. *OR Spektrum*, 19(2), pp.109–110.

Axsäter, S., 2003. Supply chain operations: serial and distribution inventory systems. In S.C. Graves and A.G. de Kok, eds. *Supply Chain Management: Design, Coordination and Operation*. Handbooks in Operations Research and Management Science. Amsterdam, North-Holland: Elsevier, pp.525–559.

Axsäter, S., 2006. *Inventory control*, New York: Springer.

Axsater, S. and Juntti, L., 1996. Comparison of echelon stock and installation stock policies for two-level inventory systems. *International Journal of Production Economics*, 45(1-3), pp.303–310.

Axsäter, S. and Juntti, L., 1997. Comparison of echelon stock and installation stock policies with policy adjusted order quantities. *International Journal of Production Economics*, 48(1), pp.1–6.

Axsäter, S. and Rosling, K., 1993. Installation vs. echelon stock policies for multilevel inventory control. *Management Science*, 39(10), pp.1274–1280.

Babai, M.Z., 2005. *Politiques de pilotage de flux dans les chaînes logistiques : impact de l'utilisation des prévisions sur la gestion de stocks*. Thesis (PhD). Châtenay-Malabry, France: Laboratoire Genie Industriel, Ecole Centrale Paris.

Billington, C., Callioni, G., Crane, B. et al., 2004. Accelerating the profitability of Hewlett-Packard's supply chains. *Interfaces*, 34(1), pp.59–72.

Bodt, M.A.D. and Graves, S.C., 1985. Continuous-review policies for a multi-echelon inventory problem with stochastic demand. *Management Science*, 31(10), pp.1286–1299.

Bossert, J.M. and Willems, S.P., 2007. A periodic-review modeling approach for guaranteed service supply chains. *Interfaces*, 37(5), pp.420–436.

Cachon, G.P., 2003. Supply chain coordination with contracts. In S. C. Graves & A. G. de Kok, eds. *Supply Chain Management: Design, Coordination and Operation*. Handbooks in Operations Research and Management Science. Amsterdam, North-Holland: Elsevier, pp.227–339.

Chen, F., 2000. Optimal policies for multi-echelon inventory problems with batch ordering. *Operations Research*, 48(3), pp.376–389.

Chen, F. and Zheng, Y.S., 1994a. Evaluating echelon stock (R, nQ) policies in serial production/inventory systems with stochastic demand. *Management Science*, 40(10), pp.1262–1275.

Chen, F. and Zheng, Y.S., 1994b. Lower bounds for multi-echelon stochastic inventory systems. *Management Science*, 40(11), pp.1426–1443.

Chen, F. and Zheng, Y.S., 1998. Near-optimal echelon-stock (R, nQ) policies in multistage serial systems. *Operations Research*, 46(4), pp.592–602.

Chu, L.Y. and Shen, Z.J.M., 2010. A Power-of-two ordering policy for one-warehouse multiretailer systems with stochastic demand. *Operations Research*, 58(2), pp.492–502.

Clark, A.J. and Scarf, H., 1960. Optimal policies for a multi-echelon inventory problem. *Management Science*, 6(4), pp.475–490.

De Kok, A.G. and Fransoo, J.C., 2003. Planning supply chain operations: definition and comparison of planning concepts. In S.C. Graves and A.G. de Kok, eds. *Supply Chain Management: Design, Coordination and Operation*. Handbooks in Operations Research and Management Science. Amsterdam, North-Holland: Elsevier, pp.597–675.

De Kok, A.G., Janssen, F., Van Doremalen, J., Van Wachem, E., Clerkx, M. and Peeters, W., 2005. Philips electronics synchronizes its supply chain to end the bullwhip effect. *Interfaces*, 35(1), pp.37–48.

De Kok, A.G., Lagodimos, A.G. and Siedel, H.P., 1994. *Stock allocation in a two-echelon distribution network under service contraints.* Research Report. The Netherlands: Eindhoven University of Technology.

Diks, E.B. and De Kok, A.G., 1998. Optimal control of a divergent multi-echelon inventory system. *European Journal of Operational Research*, 111(1), pp.75–97.

Diks, E.B. and De Kok, A.G., 1999. Computational results for the control of a divergent N-echelon inventory system. *International Journal of Production Economics*, 59(1–3), pp.327–336.

Diks, E.B., De Kok, A.G. and Lagodimos, A.G., 1996. Multi-echelon systems: A service measure perspective. *European Journal of Operational Research*, 95(2), pp.241–263.

Doğru, M.K., De Kok, A.G. and Van Houtum, G.J., 2009. A numerical study on the effect of the balance assumption in one-warehouse multi-retailer inventory systems. *Flexible Services and Manufacturing Journal*, 21(3-4), pp.114–147.

Egri, P., 2012. Safety stock placement in non-cooperative supply chains. In *Proceedings of the 3rd Workshop on Artificial Intelligence and Logistics*. AILog'12. Montpellier, France, pp.31–36.

Eppen, G. and Schrage, L., 1981. Centralized ordering policies in a multi-warehouse system with lead times and random demand. In L.B. Schwarz, ed. *Multi-level production-inventory control systems: theory and practice*. North-Holland, pp.51–67.

Eruguz, A.S., Jemai, Z., Sahin, E. and Dallery, Y., 2012. A review of the guaranteed-service model for multi-echelon inventory systems. In *Proceedings of the 14th IFAC Symposium on Information Control Problems in Manufacturing*. INCOM 2012. Bucharest, Romania, pp.1439–1444. Available at: http://dx.doi.org/10.3182/20120523-3-RO-2023.00090.

Eruguz, A.S., Jemai, Z., Sahin, E. and Dallery, Y., 2013a. Applying the guaranteed-service model to a decentralized supply chain: Impact on the cycle-service-level. In *Proceedings of the 5th International Conference on Industrial Engineering and Systems Management*. IESM'2013. Rabat, Morocco, pp.361–366.

Eruguz, A.S., Jemai, Z., Sahin, E. and Dallery, Y., 2013b. *Customer service level evaluation in decentralized guaranteed-service supply chains*. Presented at the Informs MSOM Conference, July 29, 2013, Fontainebleau, France.

Eruguz, A.S., Jemai, Z., Sahin, E. and Dallery, Y., 2013c. Cycle-service-level in guaranteed-service supply chains. In *Proceedings of the 5th International Conference on Modeling, Simulation and Applied Optimization*. ICMSAO'13. Hammamet, Tunisia, pp.1–6. Available at: http://dx.doi.org/10.1109/ICMSAO.2013.6552572.

Eruguz, A.S., Jemai, Z., Sahin, E. and Dallery, Y., 2013d. *Quantifying the impacts of the guaranteed-service model assumptions*. Presented at the 11th ISIR Summer School on Research Trends in Inventory Management and Modeling, August 20, 2013, Neuchatel, Switzerland.

Eruguz, A.S., Jemai, Z., Sahin, E. and Dallery, Y., 2014. Optimising reorder intervals and order-up-to levels in guaranteed service supply chains. *International Journal of Production Research*, 52(1), pp.149–164.

Ettl, M., Feigin, G.E, Lin, G.Y. and Yao, D.D., 2000. A supply network model with base-stock control and service requirements. *Operations Research*, 48(2), pp.216–232.

Farasyn, I., Humair, S., Kahn, J.I. et al., 2011. Inventory Optimization at Procter & Gamble: Achieving Real Benefits Through User Adoption of Inventory Tools. *Interfaces*, 41(1), pp.66 –78.

Federgruen, A. and Zipkin, P., 1984a. Approximations of dynamic, multilocation production and inventory problems. *Management Science*, 30(1), pp.69–84.

Federgruen, A. and Zipkin, P., 1984b. Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4), pp.818–836.

Funaki, K., 2012. Strategic safety stock placement in supply chain design with due-date based demand. *International Journal of Production Economics*, 135, pp.4–13.

Gallego, G., Özer, Ö. and Zipkin, P., 2007. Bounds, heuristics, and approximations for distribution Systems. *Operations Research*, 55(3), pp.503–517.

Glasserman, P. and Tayur, S., 1995. Sensitivity analysis for base-stock levels in multiechelon production-inventory systems. *Management Science*, 41(2), pp.263–281.

Graves, S.C., 1996. A multiechelon inventory model with fixed replenishment intervals. *Management Science*, 42(1), pp.1–18.

Graves, S.C. and Willems, S.P., 1998. *Optimizing strategic safety stock placement in supply chains*. Working Paper. Massachusetts Institute of Technology, Sloan School of Management.

Graves, S.C. and Willems, S.P., 2000. Optimizing strategic safety stock placement in supply chains. *Manufacturing & Service Operations Management*, 2(1), pp.68–83.

Graves, S.C. and Willems, S.P., 2003. Supply chain design: Safety stock placement and supply chain configuration. In S.C. Graves and A.G. de Kok, eds. *Supply Chain Management: Design, Coordination and Operation*. Handbooks in Operations Research and Management Science. Amsterdam, North-Holland: Elsevier, pp.95–132.

Graves, S.C. and Willems, S.P., 2005. Optimizing the supply chain configuration for new products. *Management Science*, 51(8), pp.1165–1180.

Graves, S.C. and Willems, S.P., 2008. Strategic inventory placement in supply chains: Nonstationary demand. *Manufacturing & Service Operations Management*, 10(2), pp.278–287.

Hadley, H.G. and Whitin, T.M., 1963. *Analysis of inventory systems*, Englewood Cliffs, New Jersey: Prentice Hall.

Hax, A.C. and Candea, D.I., 1984. *Production and inventory management*, Englewood Cliffs, New Jersey: Prentice Hall.

Hennet, J.-C., 2009. A globally optimal local inventory control policy for multistage supply chains. *International Journal of Production Research*, 47(2), pp.435–453.

Hennet, J.-C. and Arda, Y., 2008. Supply chain coordination: A game-theory approach. *Engineering Applications of Artificial Intelligence*, 21(3), pp.399–405.

Huh, W.T. and Janakiraman, G., 2010. On the optimal policy structure in serial inventory systems with lost sales. *Operations Research*, 58(2), pp.486–491.

Humair, S., Ruark, J.D., Tomlin, B. and Willems, S.P., 2013. Incorporating stochastic lead times into the guaranteed service model of safety stock optimization. *Interfaces*, 43(5), pp.421-434.

Humair, S. and Willems, S.P., 2006. Optimizing strategic safety stock placement in supply chains with clusters of commonality. *Operations Research*, 54(4), pp.725–742.

Humair, S. and Willems, S.P., 2011. Technical note—Optimizing strategic safety stock placement in general acyclic networks. *Operations Research*, 59(3), pp.781 –787.

Hwarng, H.B., Chong, C.S.P., Xie, N. and Burgess, T.F., 2005. Modelling a complex supply chain: understanding the effect of simplified assumptions. *International Journal of Production Research*, 43(13), pp.2829–2872.

Inderfurth, K., 1991. Safety stock optimization in multi-stage inventory systems. *International Journal of Production Economics*, 24(1-2), pp.103–113.

Inderfurth, K., 1992. Multistage safety stock optimization with dynamic-programming. *OR Spektrum*, 14(1), pp.19–32.

Inderfurth, K., 1993. Valuation of leadtime reduction in multi-stage production systems. In P.D.G. Fandel, P.D.T. Gulledge, and D. A. Jones, eds. *Operations Research in Production Planning and Control*. Springer Berlin Heidelberg, pp.413–427.

Inderfurth, K. and Minner, S., 1998. Safety stocks in multi-stage inventory systems under different service measures. *European Journal of Operational Research*, 106(1), pp.57–73.

Jackson, P.L., 1988. Stock allocation in a two-echelon distribution system or "what to do until your ship comes in". *Management Science*, 34(7), pp.880–895.

Jemai, Z. and Karaesmen, F., 2007. Decentralized inventory control in a two-stage capacitated supply chain. *IIE Transactions*, 39, pp.501–512.

Johnson, M.E., Lee, H.L., Davis, T. and Hall, R., 1995. Expressions for item fill rates in periodic inventory systems. *Naval Research Logistics*, 42(1), pp.57–80.

Klosterhalfen, S., 2010. *Multiple sourcing in single- and multi-echelon inventory systems*. Thesis (PhD). Mannheim, Germany: University of Mannheim.

Klosterhalfen, S. and Minner, S., 2007. Comparison of stochastic- and guaranteed-service approaches to safety stock optimization in supply chains. In K.H. Waldmann and U.M. Stocker, eds. *Operations Research Proceedings 2006*. Springer Berlin Heidelberg, pp.485–490.

Klosterhalfen, S. and Minner, S., 2010. Safety stock optimisation in distribution systems: a comparison of two competing approaches. *International Journal of Logistics Research and Applications*, 13(2), pp.99–120.

Klosterhalfen, S., Dittmar, D. and Minner, S., 2013. An integrated guaranteed- and stochastic-service approach to inventory optimization in supply chains. *European Journal of Operational Research*, 231(1), pp.109–119.

Langenhoff, L.J.G. and Zijm, W.H.M., 1990. An analytical theory of multi-echelon production/distribution systems. *Statistica Neerlandica*, 44(3), pp.149–174.

Lee, H.L. and Billington, C., 1993. Material management in decentralized supply chains. *Operations Research*, 41(5), pp.835–847.

Lesnaia, E., 2004. *Optimizing safety stock placement in general network supply chains*. Thesis (PhD). Massachusetts Institute of Technology, Sloan School of Management.

Li, H. and Jiang, D., 2012. New model and heuristics for safety stock placement in general acyclic supply chain networks. *Computers & Operations Research*, 39(7), pp.1333–1344.

Li, H. and Womer, K., 2008. Modeling the supply chain configuration problem with resource constraints. *International Journal of Project Management*, 26(6), pp.646–654.

Li, P., 2013. *Optimization of (R, Q) policies for multi-echelon inventory systems with guaranteed service*. Thesis (PhD). Troyes, France: Université de Technologie de Troyes.

Li, P. and Chen, H., 2012. Optimal (R, nQ) policies for serial inventory systems with guaranteed service. In D. Klatte, H.J. Lüthi and K. Schmedders, eds. *Operations Research Proceedings 2011*. Springer Berlin Heidelberg, pp.439–444.

Li, P., Chen, H. and Che, A., 2013. Optimal batch ordering policies for assembly systems with guaranteed service. *International Journal of Production Research*, 51(20), pp.6275–6293.

Lin, G., Ettl, M., Buckley, S. et al., 2000. Extended-enterprise supply-chain management at IBM personal systems group and other divisions. *Interfaces*, 30, pp.7–25.

Magnanti, T., Shen, Z.J.M, Shu, J., Simchi-Levi, D. and Teo, C., 2006. Inventory placement in acyclic supply chain networks. *Operations Research Letters*, 34(2), pp.228–238.

Marklund, J. and Rosling, K., 2012. Lower bounds and heuristics for supply chain stock allocation. *Operations Research*, 60(1), pp.92–105.

Maxwell, W.L. and Muckstadt, J.A., 1985. Establishing consistent and realistic reorder intervals in production-distribution systems. *Operations Research*, 33(6), pp.1316–1341.

Minner, S., 1997. Dynamic programming algorithms for multi-stage safety stock optimization. *OR Spektrum*, 19(4), pp.261–271.

Minner, S., 2000. *Strategic safety stocks in supply chains*, Berlin: Springer.

Minner, S., 2001. Strategic safety stocks in reverse logistics supply chains. *International Journal of Production Economics*, 71(1-3), pp.417–428.

Muckstadt, J.A. and Roundy, R.O., 1993. Chapter 2 Analysis of multistage production systems. In *Logistics of Production and Inventory*. Amsterdam: Elsevier, pp.59–131.

Muckstadt, J.A. and Sapra, A., 2010. *Principles of inventory management*, New York: Springer.

Neale, J.J. and Willems, S.P., 2009. Managing inventory in supply chains with nonstationary demand. *Interfaces*, 39, pp.388–399.

Nepal, B., Monplaisir, L. and Famuyiwa, O., 2011. A multi-objective supply chain configuration model for new products. *International Journal of Production Research*, 49(23), pp.7107–7134.

Parker, R.P. and Kapuscinski, R., 2004. Optimal policies for a capacitated two-echelon inventory system. *Operations Research*, 52(5), pp.739–755.

Rambau, J. and Schade, K., 2010. The stochastic guaranteed service model with recourse for multi-echelon warehouse management. *Electronic Notes in Discrete Mathematics*, 36, pp.783–790.

Rosling, K., 1989. Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4), pp.565–579.

Roundy, R., 1985a. *94%-Effective lot-sizing in multi-stage assembly systems*. Technical Report. Cornell University, School of Operations Research and Industrial Engineering.

Roundy, R., 1985b. 98%-Effective integer-ratio lot-sizing for one-warehouse multi-retailer systems. *Management Science*, 31(11), pp.1416–1430.

Roundy, R., 1986. A 98%-Effective lot-sizing rule for a multi-product, multi-stage production / inventory system. *Mathematics of Operations Research*, 11(4), pp.699–727.

Schneider, H., 1981. Effect of service-levels on order-points or order-levels in inventory models. *International Journal of Production Research*, 19(6), pp.615–631.

Schoenmeyr, T., 2008. *Strategic inventory placement in multi-echelon supply chains : three essays*. Thesis (PhD). Massachusetts Institute of Technology, Sloan School of Management.

Schoenmeyr, T. and Graves, S.C., 2009. Strategic safety stocks in supply chains with evolving forecasts. *Manufacturing & Service Operations Management*, 11(4), pp.657–673.

Schwarz, L.B., 1973. A simple continuous review deterministic one-warehouse N-retailer inventory problem. *Management Science*, 19(5), pp.555 –566.

Shang, K.H., 2008. Note: A simple heuristic for serial inventory systems with fixed order costs. *Operations Research*, 56(4), pp.1039–1043.

Shang, K.H. and Zhou, S.X., 2010. *Optimizing replenishment intervals for two-echelon distribution systems with stochastic demand.* Working Paper. Fuqua School of Business, Duke University.

Sherbrooke, C.C., 1968. Metric: A multi-echelon technique for recoverable item control. *Operations Research*, 16(1), pp.122–141.

Shi, J. and Zhao, Y., 2010. Technical note: Some structural results on acyclic supply chains. *Naval Research Logistics*, 57(6), pp.605–613.

Shu, J. and Karimi, I.A., 2009. Efficient heuristics for inventory placement in acyclic networks. *Computers and Operations Research*, 36(11), pp.2899–2904.

Silver, E.A. and Bischak, D.P., 2011. The exact fill rate in a periodic review base stock system under normally distributed demand. *Omega*, 39(3), pp.346–349.

Silver, E.A., Pyke, D.F. and Peterson, R., 1998. *Inventory management and production planning and scheduling*, Hoboken, New Jersey: John Wiley & Sons.

Simchi-Levi, D. and Zhao, Y., 2005. Safety stock positioning in supply chains with stochastic lead times. *Manufacturing & Service Operations Management*, 7(4), pp.295–318.

Simchi-Levi, D. and Zhao, Y., 2012. Performance evaluation of stochastic multi-echelon inventory systems: A survey. *Advances in Operations Research*, 2012, pp.1–34.

Simpson, K.F., 1958. In-process inventories. *Operations Research*, 6(6), pp.863–873.

Sitompul, C., Aghezzaf, E.H., Dullaert, W. and Van Landeghem, H., 2008. Safety stock placement problem in capacitated supply chains. *International Journal of Production Research*, 46(17), pp.4709–4727.

Tian, F., Willems, S.P. and Kempf, K.G., 2011. An iterative approach to item-level tactical production and inventory planning. *International Journal of Production Economics*, 133(1), pp.439–450.

Van der Heijden, M.C., Diks, E.B. and De Kok, A.G., 1997. Stock allocation in general multi-echelon distribution systems with (R, S) order-up-to-policies. *International Journal of Production Economics*, 49(2), pp.157–174.

Van Donselaar, K.H., 1989. *Material coordination under uncertainty*. Thesis (PhD). Eindhoven: Technische Universiteit Eindhoven.

Van Houtum, G.J., 2006. Multi-echelon production/inventory systems: Optimal policies, heuristics, and algorithms. In *Tutorials in Operations Research 2006*. Hanover, MD, USA: INFORMS, pp.163–199.

Van Houtum, G.J., Inderfurth, K. and Zijm, W.H.M., 1996. Materials coordination in stochastic multi-echelon systems. *European Journal of Operational Research*, 95(1), pp.1–23.

Van Houtum, G.J., Scheller-Wolf, A. and Yi, J., 2007. Optimal control of serial inventory systems with fixed replenishment intervals. *Operations Research*, 55(4), pp.674–687.

Van Houtum, G.J. and Zijm, W.H.M., 1991. Computational procedures for stochastic multi-echelon production systems. *International Journal of Production Economics*, 23(1-3), pp.223–237.

Van Houtum, G.J. and Zijm, W.H.M., 2000. On the relationship between cost and service models for general inventory systems. *Statistica Neerlandica*, 54(2), pp.127–147.

Wieland, B., Mastrantonio, P., Willems, S.P. and Kempf, K.G., 2012. Optimizing inventory levels within Intel's channel supply demand operations. *Interfaces*, 42(6), pp.517–527.

Willems, S.P., 2008. Data set—Real-world multiechelon supply chains used for inventory optimization. *Manufacturing & Service Operations Management*, 10(1), pp.19–23.

Yao, M.J. and Wang, Y.C., 2006. A new algorithm for one-warehouse multi-retailer systems under stationary nested policy. *Optimization Methods & Software*, 21(1), pp.41–56.

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), pp.7802–7817.

You, F. and Grossmann, I.E., 2010. Integrated multi-echelon supply chain design with inventories under uncertainty: MINLP models, computational strategies. *AIChE Journal*, 56(2), pp.419–440.

You, F. and Grossmann, I.E., 2011. Stochastic inventory management for tactical process planning under uncertainties: MINLP models and algorithms. *Aiche Journal*, 57(5), pp.1250–1277.

Zhang, Jiang and Zhang, Jun, 2007. Fill rate of single-stage general periodic review inventory systems. *Operations Research Letters*, 35(4), pp.503–509.

Zhao, Y., 2008. Evaluation and optimization of installation base-stock policies in supply chains with compound poisson demand. *Operations Research*, 56(2), pp.437–452.

Zipkin, P.H., 2000. *Foundations of inventory management*, Singapore: McGraw-Hill.