2008

Multiobjective Coordination Models For Maintenance And Service Parts Inventory Planning And Control

Oscar Martinez
University of Central Florida

MULTIOBJECTIVE COORDINATION MODELS FOR MAINTENANCE AND SERVICE PARTS INVENTORY PLANNING AND CONTROL

by

OSCAR E. MARTINEZ
B.S., Pontificia Universidad Javeriana, 2001
M.S., University of Central Florida, 2005

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Industrial Engineering and Management Systems
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2008

Major Professor: Christopher D. Geiger

ABSTRACT

In many equipment-intensive organizations in the manufacturing, service and particularly the defense sectors, service parts inventories constitute a significant source of tactical and operational costs and consume a significant portion of capital investment. For instance, the Defense Logistics Agency manages about 4 million consumable service parts and provides about 93% of all consumable service parts used by the military services. These items required about US$1.9 billion over the fiscal years 1999-2002. During the same time, the US General Accountability Office discovered that, in the United States Navy, there were about 3.7 billion ship and submarine parts that were not needed. The Federal Aviation Administration says that 26 million aircraft parts are changed each year. In 2002, the holding cost of service parts for the aviation industry was estimated to be US$50 billion. The US Army Institute of Land Warfare reports that, at the beginning of the 2003 fiscal year, prior to Operation Iraqi Freedom the aviation service parts alone was in excess of US$1 billion. This situation makes the management of these items a very critical tactical and strategic issue that is worthy of further study. The key challenge is to maintain high equipment availability with low service cost (*e.g.*, holding, warehousing, transportation, technicians, overhead, *etc.*). For instance, despite reporting US$10.5 billion in appropriations spent on purchasing service parts in 2000, the United States Air Force (USAF) continues to report shortages of service parts. The USAF estimates that, if the investment on service parts decreases to about US$5.3 billion, weapons systems availability would range from 73 to 100 percent. Thus, better management of service parts inventories should create opportunities for cost savings caused by the efficient management of these inventories. Unfortunately, service parts belong to a class of inventory that continually makes them difficult

to manage. Moreover, it can be said that the general function of service parts inventories is to support maintenance actions; therefore, service parts inventory policies are highly related to the resident maintenance policies. However, the interrelationship between service parts inventory management and maintenance policies is often overlooked, both in practice and in the academic literature, when it comes to optimizing maintenance and service parts inventory policies. Hence, there exists a great divide between maintenance and service parts inventory theory and practice.

This research investigation specifically considers the aspect of joint maintenance and service part inventory optimization. We decompose the joint maintenance and service part inventory optimization problem into the supplier's problem and the customer's problem. Long-run expected cost functions for each problem that include the most common maintenance cost parameters and service parts inventory cost parameters are presented. Computational experiments are conducted for a single-supplier two-echelon service parts supply chain configuration varying the number of customers in the network. Lateral transshipments (LTs) of service parts between customers are not allowed. For this configuration, we optimize the cost functions using a traditional, or decoupled, approach, where each supply chain entity optimizes its cost individually, and a joint approach, where the cost objectives of both the supplier and customers are optimized simultaneously. We show that the multiple objective optimization approach outperforms the traditional decoupled optimization approach by generating lower system-wide supply chain network costs. The model formulations are extended by relaxing the assumption of no LTs between customers in the supply chain network. Similar to those for the no LTs configuration, the results for the LTs configuration show that the multiobjective optimization outperforms the decoupled optimization in terms of system-wide cost. Hence, it is economically beneficial to jointly consider all parties within the supply network. Further, we compare the model configurations – LTs versus no LTs, and we show that using LTs improves the overall savings of the system. It is observed that the improvement is mostly derived from reduced shortage costs since the equipment downtime is reduced due to the proximity of the supply.

The models and results of this research have significant practical implications as they can be used to assist decision-makers to determine when and where to pre-position parts inventories to maximize equipment availability. Furthermore, these models can assist in the preparation of the terms of long-term service agreements and maintenance contracts between original equipment manufacturers and their customers (*i.e.*, equipment owners and/or operators), including determining the equitable allocation of all system-wide cost savings under the agreement.

I dedicate this dissertation to my family and the people that in some way or another inspired me and gave me the courage to complete this journey. First of all, to the memory of my father, Jaime Enrique Martinez Deluque, who always emphasized the importance of education and was a role model for hard work, dedication and honesty. Also, to the memory of my grandmother, Deyanira Deluque, for being an inspiration for her tenacity and for her love for life. To my mother, Lucila Diaz, for all her love, patience, words of encouragement, and because she made me the person I am now. Finally, to my sisters, Angela Patricia and Marian Karina who have always been with me in good and bad times and are very special.

ACKNOWLEDGMENTS

This dissertation could have not been written without the advice and dedication of Dr. Christopher D. Geiger, my advisor and dissertation committee chair. Special thanks to him for his countless hours invested in this research and his patience throughout the entire process. Without his constant feedback, it would have been difficult to maintain the quality that this project required. Thanks to Dr. Charles Reilly, Dr. Stephen Goodman, Dr. Robert Hoekstra and Dr. Dima Nazzal for their participation on my dissertation committee, and their valuable feedback.

Thanks are owed to Dr. Ronald Eaglin for his financial support during most of my graduate studies at the University of Central Florida. I also want to thank Dr. José Sepúlveda and Dr. Luis Rabelo. They both opened the door for me to enter graduate school at the University of Central Florida and taught me how to succeed in it.

Finally, I must acknowledge many friends, colleagues, students, and professors who assisted, advised, and supported my research and writing efforts over the years. Especially, I express my gratitude and deep appreciation to Mehmet Onal and Dr. Sami Spahi whose friendship and knowledge have supported, enlightened and entertained me over the many years of our friendship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1. The State of Service Parts Inventory Management

Equipment-intensive organizations including service, military and manufacturing companies are constantly challenged by their environment. These organizations must operate their business with high system availability, *e.g.*, utility companies and commercial airlines. To ensure continuity of operations, an ample supply of service parts (sometimes referred to as spare parts) should be maintained. The failure of a component can cause the system to fail potentially causing severe downtime consequences. In contrast, maintaining a high inventory of service parts ties up capital and often results in exorbitant inventory costs.

For many organizations, service parts inventories consume a significant portion of their capital investment. For instance, the Defense Logistics Agency (DLA) manages about 4 million consumable service parts and provides about 93% of all consumable service parts used by the military services. These items required about US$1.9 billion investment over the fiscal years 1999-2002. On the other hand, the US General Accountability Office (GAO) discovers that, in the United States Navy, there were about 3.7 billion ship and submarine parts that were not needed. The Federal Aviation Administration (FAA) says that 26 million aircraft parts are changed each year. In 2002, the holding cost of service parts for the aviation industry is estimated to be US$50 billion (Kilpi *et al.*, 2004).

This situation makes the management of these items a very critical issue that is worthy of careful study. As indicated by Diaz and Fu (2005), in many organizations in the manufacturing, services and defense sectors, there are opportunities for cost savings by engaging in more efficient management of service parts inventories, and the trend is that it is likely to become even more critical. The key challenge is to maintain high availability with low service cost (*e.g.*, holding, warehousing, transportation, technicians, overhead, *etc.*). For instance, despite reporting US$10.5 billion in appropriations spent on purchasing service parts in 2000, the United States Air Force (USAF) continues to report shortages of service parts. The USAF estimates that, if the investment on service parts decreases to about US$5.3 billion, weapons systems availability would range from 73 to 100 percent (GAO, 2003).

Many organizations are expending enormous efforts to improve the management of service parts, especially in the defense sector, where equipment readiness is a critical aspect of military readiness. For instance, the USAF has identified more than 80 initiatives to address more than 300 deficiencies involving important aspects such as the processes that affect service parts shortages, depot maintenance, and supply of service parts (GAO, 2003). In industry, IBM, for example, has created different tools to manage their more than 50 million service parts (Cohen, 1990).

By their very nature, service parts belong to class of inventory that makes them inherently difficult to manage. The general function of service parts inventories is to assist a maintenance staff in keeping equipment in operating condition. Consequently, the policies that govern service parts inventories are different from those that govern other type of inventories such as raw material inventory, work-in-process inventory and finished goods inventory. Service parts are characterized by low demand with stochastic and frequent irregular patterns. Hence, service parts inventory management is one important area that should not be underestimated.

1.2. Multi-Echelon Service Parts Inventory Systems and Inventory Pooling

Common in industry, different levels (echelons) frequently make up a service parts supply chain, where the multi-echelon inventory system involves the existence of a hierarchy of parts stocking locations. For instance, Siemens AG Power Generation (2004) implements a three-tier service parts logistics strategy to reduce its stockpiling of service parts for its Instrumentation and Controls Services business unit that supports the other business units. The first tier requires that service parts critical to production are stored onsite at the customer. The second tier is the use of a centralized stocking location that is managed by Siemens. The third tier employs a logistic service that delivers what are considered non-critical parts when requested by customers. Siemens reports that the onsite service parts inventory can be reduced by up to 80%. The use of multi-echelon inventory models for the management of service parts inventory is primarily driven by a need to control and reduce the inventory holding costs as well as the improvement of the response time in the supply of items. However, it is the dependencies and interactions between the different echelons that complicate the service parts inventory control problem.

Further, many organizations have extended the multi-echelon configurations to share and collaborate among partnering organizations (through inventory pooling) with the goal of benefiting not only the individual partnering organizations but also the entire supply chain. Inventory pooling, which is typically an agreement between partners belonging to the same echelon to share their service parts inventories, has been shown to be an effective strategy to improve logistical performance and reduce overall inventories in multi-echelon supply chain systems. The idea of pooling inventories relies on timely lateral transshipments between pooling partners. This cooperation can also be used to improve the service levels and equipment availability while reducing the total supply chain system cost. The cost of a shipment from a pooling partner is generally much lower than the combined costs of equipment downtime and emergency part shipments from a higher echelon. In general, these situations are usually economically more attractive than the costs incurred in a regular emergency shipment from the parts supplier. Past research suggests that lateral transshipments can reduce overall equipment downtime cost and inventory cost. However, this is not an easy problem to address, as indicated by Köchel and Nieländer (2002). One of the main problems is how to define optimal order and transshipment policies in multi-echelon inventory models.

In recent years, several companies came to understand the benefits of service parts inventory pooling and have entered into such partnerships, particularly in the airline industry. For instance, the Boeing Company has a worldwide agreement called Spares Exchange Program. This agreement currently includes KLM Royal Dutch Airlines, Transavia Airlines of the Netherlands, Braathens of Norway, Polynesian Airlines, Oman Air, and Yemen Airways. Also KSSU, a maintenance consortium in Europe, includes KLM, SAS, Swissair and UTA. Atlas has grouped Air France, Alitalia, Iberia and Lufthansa. Another industry where companies realize the importance and benefits of pooling service parts inventory is the power generation industry. General Electric Energy has pooling agreements with its transformer customers. The ABB Group establishes these agreements with its customers for power generation, water and wastewater plants. Power Engineering International (2003) reports that several public power generation companies in Florida have formed a common organization to manage the collective service parts for their combustion turbines in an effort to reduce their individual stockpiles.

1.3. Long-Term Service Agreements

The immediate availability of critical service parts is a must in order to improve efficiency, productivity and safety, extend equipment life cycles and minimize equipment downtime, and in some cases, as in the power generation sector, comply with regulatory and environmental regulations. The probability of success in achieving these objectives increases significantly when the equipment is covered by a long-term service agreement (LTSA) managed by the original equipment manufacturer (OEM). In principle, an LTSA is an agreement between the owner (or operator) of the equipment (*i.e.*, the customer) and the OEM (*i.e.*, the parts supplier). These agreements, sometimes referred to as maintenance service agreements, service contracts, or maintenance contracts, are essentially risk management tools that are used quite commonly in the power generation market. For instance, General Electric Power Systems and the Power Generation Group of Siemens AG offer and maintain these for their customers. Levallois-Perret, France-headquartered Alstom, currently offers operation and maintenance contracts to its power customers.

The spectrum of LTSAs has expanded over the years, and these agreements can range from simple purchase orders to service agreements that ensure the timely delivery of critical service parts to agreements that allow partial or perhaps full facility operation and maintenance by the OEM. These agreements ensure a desired level of responsibility of the OEM to a customer, from the provisioning of a single service or product to the responsibility for operating and maintaining the entire power plant (if requested), including power availability guarantees. The benefit to the customer is that, with the increased involvement and responsibility of the OEM, the risk of the owner or operator of the power plant is decreased. Lehmann (2006) of Siemens Power Generation lists some of the major benefits of LTSAs to the customer as:

- discounted OEM parts and services;
- extended warranties, which means more than the normal service or product guarantee; and
- remote performance monitoring and diagnostic services.

In short, LTSAs can significantly reduce equipment repair time consequently reducing planned and unplanned production outages.

The provisioning of critical service parts is often a major component of LTSAs, and the availability and access of parts are among the major factors leading to a reduction of downtime when a breakdown occurs. If the required parts are not available in stock at the customer, the equipment cannot be repaired until the part is ordered and shipped from the supplier. Firms have realized the importance of appropriate storage and timely delivery of service parts for their customers. Another important aspect of the LTSAs is the pooling of inventories. So the supplier, in certain situations, can make use of lateral transshipments between customers to cover emergency orders.

A major challenge of LTSAs is negotiating the most appropriate terms under which the customer and the part supplier will partner as well as the responsibilities of each party. The terms of the LTSA are negotiated so that the operational and strategic objectives of all parties are achieved. The two primary parties involved in LTSAs that have a major interest in them are the customers and the parts supplier, and the successful, effective agreement should be crafted from the viewpoint of both parties. The challenge though is to strike a balance between multiple and conflicting operational objectives of the parties, *i.e.*, maintain high system availability with low cost (*e.g.*, inventory holding, transportation/delivery, production downtime, labor, overhead, *etc.*).

1.4. Multiobjective Optimization of Service Parts Inventory Supply Systems

Several authors have addressed the problem of determining the optimal levels of service parts (*e.g.*, Rudi *et al.*, 2001; Marseguerra *et al.*, 2004; Wong *et al.*, 2006). However, some of the existing modeling approaches and solution methods in the literature use simplified models whose applicability to real-world multi-echelon inventory systems may be questionable. Some other authors (*e.g.*, Köchel *et al*., 2002; Axsäter, 2003; Zamperini *et al*., 2005) use simulation to the posterior optimization of the results for scenarios with one customer or one independent system. However, to date, no work has been done regarding the simultaneous optimization of the inventory and maintenance policies of the members (*i.e.*, customers and suppliers) of an inventory supply chain of service parts.

In this research, we propose an approach for the joint optimization of different independent systems that are members of a service parts supply chain. Each party in this multi-echelon supply chain system has its own objectives (and constraints) regarding optimizing service parts inventories and maintenance policies, and these objectives may contrast or be similar to those for the other entities in the supply chain system. This type of system clearly represents a multiobjective optimization situation. Consider a general multiobjective optimization problem (MOOP) with vector $\mathbf{x}$ of p policy decision variables (*i.e.*, x_i where $i = 1, \ldots, p$) and n objectives, where $n > 1$. The problem can be generally expressed as follows:

$$\min\ (\max)\ z_i = f_i(\mathbf{x}),\ i = 1, \ldots, n, \tag{1.1}$$

where a solution $\mathbf{x}$ is a p-dimensional vector of decision variables in the decision space that are continuous or discrete, or both. Eq. 1.1 is subject to m inequality constraints

$$g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, m,$$

and k equality constraints

$$h_l(\mathbf{x}) = 0,\ l = 1, \ldots, k.$$

Solutions that dominate the others but do not dominate themselves are called nondominated solutions. A solution $\mathbf{x}$ is said to dominate a solution $\mathbf{y}$ if and only if

$$f_i(\mathbf{x}) \leq f_i(\mathbf{y}) \text{ for all } i \in \{1, 2, \ldots, n\}, \text{ and}$$

$$f_i(\mathbf{x}) < f_i(\mathbf{y}) \text{ for at least one } i \in \{1, 2, \ldots, n\}.$$

In other words, there is at least one solution **x** that is better than a solution **y** while the remaining solutions are either equal to or worse than the others. The solution **x** is said to be Pareto optimal if no other solution can be found to dominate **x** using this definition of solution dominance. We note here that this definition can be applied to a maximization or minimization problem with no loss of generality.

Past attempts to solve multiobjective joint inventory-maintenance optimization problems have typically led researchers and practitioners to convert the MOOP into a single objective optimization problem. This approach ultimately transforms the original multiple objective formulation into a single objective optimization problem with a single, unique solution (decision). Some researchers and practitioners choose to model one objective, say as a cost (or profit) objective, and then represent other objectives as constraints. Others choose to represent the multiplicity of objectives as a single composite weighted objective function using a vector of user-defined weights in order to produce a tractable problem. Several drawbacks of using such traditional methods include:

1) the need for appropriate selection of the weight vector;
2) missing some of the optimal solutions in the nonconvex objective space; and
3) the inability to easily homogenize different quantities, such as cost, quality and time, to a common unit of measure.

These and other known drawbacks to traditional multiobjective optimization approaches have motivated researchers and practitioners to seek alternative techniques to find a set of Pareto optimal solutions rather than just a single solution. In general, in multiobjective optimization problems, the solutions are not uniquely determined. In fact, in these problems, especially where two or more objectives conflict, there usually exist many solutions that satisfy all relevant objectives, where the most desirable solution, or at least the best-compromised solution, is selected from among them. In practice, the availability of a set of Pareto optima provides the stakeholder(s) with the opportunity to consider several possible decision alternatives, which is generally preferred by decision-makers. In this case, the decision-maker can select a solution in accordance to his or her preferences or constraints. Therefore, the development of multiobjective optimization solution procedures using nondomination-based methods is preferred.

1.5. Motivation of This Research

Important aspects regarding the optimization of service parts inventory and maintenance policies in a multi-echelon supply chain system have been discussed. Two aspects motivate this research. First, the simultaneous optimization of inventories and maintenance has not been largely addressed. Second, it is clear that there is an important relationship between service parts management activities and maintenance activities since one supports the other one. However, these two areas have mainly been considered separately. For instance, the maintenance literature often assumes an infinite supply of service parts and inventory personnel do not consider maintenance needs. For this reason, developing effective strategies considering both as a joint activity is needed.

In addition to the joint optimization of service parts inventory and maintenance, the simultaneous optimization of service parts inventory and maintenance from both the customer and supplier perspectives has not traditionally been the focus in research or in practice. A multiobjective inventory supply chain system must consider each member as a separate entity. In this case, since each member has its own goals, the system should find a balance between those and determine what is better for all parties involved. In other words, the decision-maker should be able to determine a range of solutions that optimizes the system and then find the most satisfying solution for all supply system members.

1.6. Objectives of This Research

Several models have been developed and many solution methods have been proposed by the research community, but none of them fully include the two main aspects of this research. In addition, the advertised success of these models in practice has not yet been fully realized (Kranenburg and van Houtum, 2004). A reason might be that many assumptions underlie these models, resulting in some essential and realistic properties not captured in the current models. However, if these assumptions are relaxed, it is not easy or even possible to get an optimal or even near-optimal solution. Since the general problem is quite difficult, we start by analyzing a somewhat simple version of the problem in order to obtain insights into a more complicated version of the problem. Then, a primary constraint of the simpler version of the problem is relaxed to more resemble a real-world situation.

Objective 1: Maintenance and Service Parts Inventory Policy Optimization for a Two-Echelon n-Customer and Single-Supplier Supply Chain System

The first research objective investigates a service part supply system that consists of a single supplier and multiple (*n*) customers in a two-echelon supply chain. A single part supply chain where the parts are non-repairable (*i.e.*, replacement) is considered. This is similar to repairable items whose repair facilities are off-site and there is a supply of repaired parts ready to be shipped. In this supply chain scenario, the service parts supplier has one objective: to minimize that includes inventory and production costs. The supplier must decide when to produce a part for each customer. Each customer desires to minimize their total inventory and maintenance costs by deciding when to order a replacement part and when to replace a failed or non-failed part. In this scenario, no lateral transshipments occur between the customers.

Objective 2: Maintenance and Service Parts Inventory Policy Optimization for a Two-Echelon n-Customer and Single-Supplier Supply Chain System with Lateral Transshipments

Under this research objective, lateral transshipments between customer locations are allowed. This supply system configuration is similar to that in Objective 1; however, service parts can move between customers. In this problem, the supplier makes the same decisions as in the previous configuration. On the other hand, each customer must also decide when to request a part from a partnering customer in the supply chain and to which customer they submit the part request.

1.7. Contributions of This Research Investigation

This research investigation has significant benefit and broad impacts. First, a contribution of this investigation is to strengthen the foundation for the optimization of multiple objectives, in particular, regarding inventory and maintenance systems. Currently, researchers have focused on single objective optimization, ignoring the advantages of simultaneous optimization, especially for decision support. Second, practitioners will benefit from integration of two areas that are usually investigated separately but are highly-related – maintenance cost modeling and service parts inventory cost modeling. In maintenance, reliability engineers usually work on the minimization of the maintenance cost ignoring the inventories and assuming that they are always available. In service parts inventories, logisticians work on minimizing costs, sometimes sacrificing equipment reliability.

Another contribution is that, from the general models created here, organizations that are entering LTSAs will have a strong basis to create their own maintenance and inventory solutions. These solutions should give those companies a competitive advantage since less capital and resources will be needed. Furthermore, recall that a major challenge of LTSAs is negotiating the most appropriate terms under which the customers and the part supplier will partner as well as the responsibilities of each party. The models and results of this research can be used to assist decision-makers in preparing the terms of LTSAs, including determining the equitable allocation of all cost savings (sacrifice) under the agreement.

1.8. Outline of This Dissertation

The remainder of this dissertation is organized as follows. In Chapter 2, we summarize the most important work reported in the literature for the areas related to this investigation such as service parts inventory modeling, maintenance modeling, joint service parts inventory optimization, and evolutionary algorithms for optimization. In Chapter 3, we develop the long-run expected cost models for a two-echelon, single supplier, n-customer configuration for joint maintenance and service parts inventory system. In this configuration, lateral transshipments of service parts between customers are not considered. We use decoupled and joint optimization and show the differences between these two approaches. In Chapter 4, the work in Chapter 3 is extended by relaxing the constraint of lateral transshipments, and new long-run expected cost models are developed and solved. The results of the two configurations (*i.e.*, with lateral transshipments and with no lateral transshipments) are compared, and the benefits of the extended model are illustrated. Finally, in Chapter 5, we present a summary of the research followed by a discussion of directions of future work.

CHAPTER 2:
LITERATURE REVIEW

2.1. Introduction

Service parts inventory modeling and optimization, and maintenance modeling and optimization are two vast bodies of research. The open literature contains a substantial amount of work in these two areas. Therefore, we review past work in these areas of research that consider multi-echelon service part inventory systems. Our review of the literature is divided into three parts. First, we begin with the previous work in service (or spare) parts inventory modeling. Next, we discuss previous work that involves maintenance modeling that supports service parts modeling. Finally, previous work in the area of multiobjective optimization, in terms of maintenance and service parts inventories, is briefly discussed.

2.2. Multi-Echelon Service Parts Inventory Modeling

In general, multi-echelon service parts inventory models have been widely researched from several different perspectives (Kennedy *et al.*, 2002). Several assumptions have been made regarding the demand arrival process, the different channels through which parts move, the number of parts in the system, *etc*. with the goal of finding optimal or near-optimal inventory policies that minimize total system cost. The majority of these models do not include the impact of the maintenance policy in their analyses.

Multi-echelon inventory models can be classified as cyclic and acyclic models. In acyclic models, demand for parts or goods flows in only one direction. For example, in the computer industry, the supplier sends the service parts to the customer. After the replacement part is installed, the failed part is no longer useful and is discarded. Cyclic models are those in which part inventory flows through the system as demanded by the different locations, so inventory is neither lost nor produced, though it may undergo some transformation. This is the case of repairable items. Once damaged, failed parts are sent to a repair facility. As soon as the parts are repaired, they are made available for future replacements (Diaz and Fu, 2004). The focus of this research is multi-echelon cyclic service parts inventory models. It is important to note here that non-repairable service parts models can be treated as cyclic models since the supplier can be considered as the repairing warehouse and the part repair time is the lead time to receive the replacement part.

One of the earliest works and most well-known in the area of multi-echelon service parts inventory systems is that of Sherbrooke (1968). In his seminal work, Sherbrooke (1968) develops the METRIC model, which is a model of a two-echelon service parts inventory supply system for the US Air Force to manage their repairable items. In this supply system, there are n bases that are supported (supplied) by a central depot. Repair is performed at the base-level and at the depot-level, depending on the required level of repair. All the facilities in the system have ample repair capacity and operate according to a continuous review (S, S-1) policy, or base stock policy. Sherbrooke (1968) considers the minimization of the total expected backorders at the bases subject to a system investment constraint. There have been several extensions to Sherbrooke's work (*e.g.*, Muckstadt, 1973; Muckstadt and Thomas, 1980; Sherbrooke, 1986; Diaz 2003). Muckstadt (1973) develops the MOD-METRIC model, which allows for multiple levels of indenture, *i.e.*, service parts requirements for an end assembly item and its components. Muckstadt and Thomas (1980) develop another extension of the METRIC model by allowing emergency replenishments in case of stockout situations at the central or local warehouses. More recently, Diaz (2003) addresses the simultaneous optimization of the service parts inventory using the METRIC model and the resources needed for maintenance.

Some studies focus on analyzing multi-echelon inventory systems with a single part. This includes the work of Moinzadeh and Lee (1986) that considers a single-part model and develops a search routine for the stocking levels. They also derive a decision rule to select an (S, S-1) versus an (Q, r) policy. Axsäter (1990a) considers a single-part, two-echelon, one-for-one replenishment (base stock) model. He develops a recursive procedure to determine average holding and shortage costs and discusses the characteristics of optimal inventory base stock levels. He concludes that the model requires more computational resources than the METRIC and the Graves approximations when only a single policy is evaluated, but requires less computational effort when the whole system is to be optimized.

Cohen (1990) develops a multi-echelon model called Optimizer to determine optimal service and service parts inventory policies of IBM. They use fill rate as the service measure and solve the problem by decomposing the model into three stages. The decomposition starts with the lowest echelon where demand occurs and passes up to the following level, and so on. They use a heuristic solution procedure in solving the service allocation problem. Each iteration of the algorithm increments to part stocking levels are selected on the basis of their marginal contributions to improving the objective function and to meeting the service constraints.

Wang *et al.* (2000) consider a two-echelon, multi-item, stochastic demand service parts system with stocking-center-dependent replenishment lead times. They characterize the system performance of the stocking policies, and they show that the random delays at the depot experienced by center replenishment orders are different from center to center. They also show that this difference can be substantial when both the planned depot stocking level and the demand rates are low. Lau *et al.* (2006) analyze a three-echelon repairable inventory system with passivation. Passivation means that, when the system is down and under repair due to a part failure, the failure rate of the other supporting parts is equal to zero. They calculate the expected backorder and the operational availability at each epoch over a given time horizon, taking into consideration non-stationary demands and the effects of passivation to capture the time-varying behavior of the objective function. In their experiments, they compare the operational availability versus the METRIC model and the time-varying demands against Monte Carlo simulation. They conclude that their model performs better than the METRIC model in the context of fast-changing business, where the assumption of constant demand-rate for inventory planning and optimization is no longer realistic.

2.2.1. Multi-Echelon Inventory Models with Lateral Transshipments

There have been other works that not only consider multi-echelon service part inventory supply systems, but also multi-echelon systems with lateral transshipments (*e.g.*, Lee, 1987; Axsäter, 1990b; Sherbrooke, 1992; Archibald *et al.*, 1997; Alfredsson and Verrijdt, 1999; Grahovac and Chakravarty, 2001; Axsäter, 2003; Caglar *et al.*, 2003; Wong, 2005b; Wong, 2006). Lee (1987) considers lateral transshipments in a two-echelon inventory system for repairable parts that employs a continuous review inventory policy. He analyzes a system that consists of one supplier (depot) and several identical locations (bases), and derives approximations of the expected values of performance measures such as the backorder level and the number of emergency shipments from the depot to bases. These approximations are used to determine the optimal stocking level in the system. He concludes that the use of emergency lateral transshipments can result in significant savings since less stock is needed at the individual bases. However, the use of emergency lateral transshipments greatly depends on the magnitudes of the cost and lead time of the lateral shipments relative to the cost and lead time of receiving the part from the depot.

Axsäter (1990b) improves Lee's model by relaxing the assumption that bases are identical. Alfredsson and Verrijdt (1999) extend the model of Axsäter (1990b) by allowing emergency shipments from not only the pooling partner, but also the central warehouses and the manufacturing plants. Sherbrooke (1992) conducts a simulation study to investigate the importance of lateral transshipments in a two-echelon depot-base system for repairable items. It is important to note that his model uses a simulation model rather than an analytical model. Archibald *et al.* (1997) consider a two-location, multi-item, multi-period, periodic review inventory system with a limited storage space for all items. They allow an emergency order for customers in which equipment downtime is very costly. They conclude that the holding cost for which individual optimal reorder policies exactly fill both depots are not unique. They also state that performing lateral transshipments early in the planning period is favored when the emergency cost plus the holding cost is large compared to the shipment cost plus the unit cost.

Alfredsson and Verrijdt (1999) and Grahovac and Chakravarty (2001) both consider non-zero lateral transshipment time with the assumptions of ample repair capacity and Poisson failure process with constant failure rate. First, they use a zero lateral transshipment time model. Then, they calculate the additional downtime by directly multiplying the average number of lateral transshipments with the average lateral transshipment time.

Wong *et al.* (2005b) propose a heuristic for a multi-item two-echelon inventory system for repairable parts with lateral transshipments and continuous review policy. They use a Lagrangian relaxation approach to obtain lower and upper bounds on the total cost. Wong *et al.*

(2005c) develop an analytical model for service parts stocking levels in a single-item, multi-hub, multi-company, repairable inventory system in which complete pooling of stock is permitted. In this configuration, the machines move from one location (hub) to another location. So, the failure can occur randomly at any location. They minimize the total system cost, which is the sum of the holding cost, downtime cost and transshipment cost. They develop an approximation method based on the METRIC model and propose a two-stage solution procedure. In the first stage, the optimal number of service parts is determined by finding the quantity that minimizes the total cost function. In the second stage, they use a heuristic to calculate the allocation of service parts that minimizes transshipment cost. With their approach, they conclude that there is economic benefit with the practice of pooling inventories and state that their methodology can be easily used to analyze a variety of problems that involve numerous hubs.

Wong *et al.* (2006) analyze a continuous review, multi-item model that considers all the repairable items in a multi-location inventory system. Lateral and emergency shipments occur in response to stockouts subject to waiting time constraints. They formulate three different initialization algorithms and then apply the steepest-descent local search method to improve the solution. They also use a Lagrangian relaxation-based approach to obtain the lower bounds on the total costs. They show that lateral transshipments provide cost savings to the system and conclude that:

- the relative cost savings of applying lateral transshipments are higher than the relative cost savings of applying the multi-item approach;
- the relative cost savings of applying lateral transshipments increase with the number of pooling members;

- the relative cost savings of applying lateral transshipments are higher when all pooling members set identical maximum waiting times; and
- the relative cost savings of applying the multi-item approach increase when the variability of inventory holding costs across items increases since it contributes to a significant portion of the total cost.

Caglar *et al.* (2003) study a two-echelon service parts inventory problem with multi-item and multi-location features. They assume that parts fail according to a Poisson process. They develop a model to minimize the system-wide inventory costs subject to response time constraints. Axsäter (2003) models a single-echelon inventory with unidirectional lateral transshipments, where he evaluates the resulting fill rates, the average stock on hand, and the backorder levels. He derives mathematical models and compares the results of the models with computer simulation. He concludes that the performance of the models reduces drastically when the transshipment cost increases and becomes comparable with the backorder cost.

2.2.2. Multi-Echelon Service Parts Inventory Models as Queuing Models

Although not directly related to this research investigation, but worth mentioning for the sake of completeness, there exists limited work where researchers view the multi-echelon supply system for repairable items as a network of queues, where the part pipelines are the queues at each repair facility (*e.g.*, Yanagi and Sasaki, 1992; Diaz and Fu, 1997; Avsar and Zijm, 2000; Sleptchenko *et al.*, 2005). Yanagi and Sasaki (1992) formulate a model based on the machine-repair queuing model. They consider non-zero, exponentially-distributed lateral transshipment times in their model and use a two-stage decomposition approach to solve the problem. They consider one repair capacity, and model the problem as a multi-dimensional Markovian problem. They state the machine-repair queuing models are less preferable than the METRIC-based models, because they are more difficult to solve due to the huge multi-dimensional state space involved. Diaz and Fu (1997) consider a model where the repair shop is modeled as a $G/G/k$ multi-class queuing system. The part flow of one item type is modeled as one class in the queuing system. Although they discuss formulations for multi-server queues, their numerical results refer to single server queues only. In addition, they discuss an alternative method that uses throughput times in the $M/G/\infty$ model so that waiting times are included. Avsar and Zijm (2000) propose an approximation for a two-echelon inventory model, where repair shops can be modeled as open Jackson queuing networks. However, their model considers only item-dedicated repair shops and is difficult to extend to multi-echelon model or models with different types of repair shops, as we consider. Sleptchenko *et al.* (2005) examine a repair system in which every item is a multi-indenture structure, which mean that each end assembly item may consist of other lower-level parts. They model repair shops using multi-class and multi-server priority queues. This adds more degrees of freedom to previous models such as capacity of the repair facilities and repair priorities. They conclude that a proper priority setting may lead to a significant reduction in the inventory investment required to attain certain target system availability (usually 10-20%). The saving opportunities are particularly high if the utilization of the repair shops is high and if the part types sharing the same repair shop have distinctly different characteristics such as price, repair time, *etc*.

2.2.3. Service Parts Inventory Models with Non-Poisson Demand

One of the fundamental assumptions of not only the METRIC model but also more recent inventory models proposed in the literature pertains to the part failure process. Typically, the failure rate is assumed to follow a fixed Poisson process. Some work has been done where this assumption is relaxed. Graves (1985) proposes the use of a negative binomial distribution that fits a two-parameter distribution to the distribution of outstanding orders for the single-base *n*-depot problem. Sokhan-Sanj *et al.* (1999) propose the use of a Hyperexponential distribution to simulate highly-variable part movements in a semiconductor manufacturing setting. By accurately capturing the actual system variability, they are able to eliminate an undesirable safety factor that has been commonly used in previous simulation studies. Zamperini (2005) proposes a multi-item model where the part failure arrival process follows a negative binomial distribution. He also assumes infinite repair capacity and deterministic repair times. He concludes that since the negative binomial distribution has two parameters, one can fit both the mean and variance of a unimodal data set, although the variance to mean ratio must be greater than one. In these and other models where a non-Poisson part arrival is assumed, the rate still remains stationary.

There are other multi-echelon models that have been proposed from different aspects than those previously mentioned. For instance, Kranenburg and van Houtum (2004) develop a multi-level spares inventory model based on customer differentiation. They develop a heuristic procedure using linear programming and product-form solutions for closed queuing networks that are known to produce exact solutions for single-item problems. This procedure generates a heuristic solution and a lower bound for the optimal cost. Koçağa and Sen (2007) develop a base stock model for service parts inventory management with demand lead times and customer rationing. This one-for-one replenishment model assumes multiple demand classes with different priorities that imply different service levels. Similar to other models, they assume that the failure rate follows a Poisson process. They conclude that customer rationing creates more savings when the arrival rate in the non-critical demand class is higher than the arrival rate in the critical demand class. Wong *et al.* (2005a) develop an analytical model to estimate the performance measures in a single-item repairable multi-echelon system composed of a supplier and several customers. They assume that pooling and lateral transshipments are permitted. They formulate the model as a multi-dimensional Markov chain problem and conclude that, by allowing lateral transshipments, the expected number of backorders in the system is considerably reduced if the cost and time of those is substantially low. Otherwise, the model can lead to suboptimal decisions.

In this section, we discuss existing work regarding service parts modeling and optimization. This area has been widely studied and modeled from different perspectives, such as via the METRIC model and its extensions and via queuing models. Many assumptions have been made in these models, so they hardly represent the real-word scenario that motivated the study. One of the most important assumptions that has been made is that of the occurrence of lateral transshipments. This is one of the most relevant aspects of this research because of its importance in service parts inventory pooling.

2.3. Maintenance Modeling

Several models have been proposed for the optimization of maintenance activities and policies (*e.g.*, Triantaphyllou *et al.*, 1997; Usher *et al.*, 1998; Murthy and Asghariza, 1999; Das and Sarkar, 1999; Cassady *et al.*, 2001; Li and Xu, 2003; Carnero, 2004). Triantaphyllou *et al.* (1997) develop a model for classifying different criteria relating to industrial maintenance, including availability, reliability, *etc*. They conclude that sensitivity analysis is a necessary tool when dealing with complex maintenance multi-criteria decision-making problems. Usher *et al.* (1998) present a model for determining an optimal maintenance and replacement schedule for a system subject to deterioration. In their model, they include the value of money over time, as well as the implications of imperfect maintenance. They use three different solution approaches – random search, branch and bound, and a genetic algorithm – to find a near-optimal solution. They conclude that the genetic algorithm provides better and faster results. The most important difference is in the number of iterations needed to find a solution. Using genetic algorithms, the number of iterations to find a good answer is up to 95% lower than the other methods. In the work by Murthy and Asghariza (1999), a model for optimal decision-making in a maintenance service operation is suggested. Using Markov chains, they determine the optimal strategy regarding pricing, customers to serve, and service channels.

Das and Sarkar (1999) consider a single-part inventory system with a Poisson failure rate and analyze the behavior of the time between failures when using preventive maintenance. They develop a mathematical probabilistic model from which several performance measures of the system are reported. The measures are: (1) average cost-benefit due to maintenance, (2) service level of product, (3) average level of inventory in the system, and (4) system productivity. They suggest that, by analyzing the costs, an optimal level of preventive maintenance can be determined. Cassady *et al.* (2001) propose a system they called selective maintenance through which the decision-maker can choose between multiple maintenance options, such as minimal repair of faulty components, replacement of faulty components and preventive maintenance. They develop an extension of previous mathematical programming models incorporating the Weibull distribution and compare different maintenance alternatives using Monte Carlo simulation. They conclude that their model is applicable to any support equipment that performs a sequence of jobs with time constraints for maintenance between jobs. They state that this model can be extended to situations in which maintenance opportunities may be limited by resources other than time or where multiple systems compete for the same maintenance resources.

Li and Xu (2003) introduce a multivariate repair model with a maintenance policy that performs imperfect repairs to failed components and coordinates random group replacements according to a predetermined timetable. They conclude that the maintenance process should be only performed on items that are more reliable when new, and that it is better to have a simultaneous replacement of components rather than independent replacements. In other words, when replacing items, it is better to also replace those items that might have been affected by the failure or malfunctioning of the repaired item.

Carnero (2004) develops a methodology where some indicators facilitate the detection of anomalies in the preventive maintenance programs. She develops a cost function to evaluate the cost of the resources needed in a predictive maintenance program. This function implies the use of several resources and activities. For that reason, she uses a genetic algorithm to find a near optimal solution for the cost function.

In summary, maintenance modeling has been studied by many researchers from different points of view, but very little has been done with the inclusion of inventory levels and more importantly inventory costs. The joint optimization of these two areas is critical since those activities depend on one another.

2.4. Joint Service Parts Inventory and Maintenance Optimization

One approach taken by many inventory researchers is to decompose the multi-echelon service parts inventory supply system by treating each location independently and then applying techniques for single location inventory models. Usually, such local optimization leads to suboptimal solutions for the multi-location supply system. Alternatively, there is limited research addressing a more accurate approach, at the cost of complexity. This alternate approach involves considering the interactions between echelons (see the references and comments in Diaz and Fu (2004)). Additionally, this approach considers the joint optimization of maintenance and service parts inventory policies. The joint optimization of both maintenance and inventory models of both the supplier and the customers has received little to no attention in the recent literature. This is the focus of this research investigation.

The traditional approach of maintenance optimization assumes that spares are always available when needed for the selection of a replacement policy that minimizes the expected maintenance costs for replacement and breakage. This approach also includes making some assumptions about the distribution of part demand, and then the selection of an ordering policy that minimizes the expected inventory costs of holding and shortage. Very little research has been done about considering the two aspects simultaneously. A notable exception is the work of Armstrong and Arkins (1996). They consider the joint optimization of service part replacement and ordering policies for a system with one component subject to random failure and room for only one spare in stock. They consider four costs – part replacement cost, part breakage cost, part holding cost and part shortage cost. In other words, they cover both the two major costs generally considered in the maintenance literature as well as the two major costs widely considered in the inventory literature. They derive a single cost function of both inventory-related and replacement-related costs based on the same scenario for a single company, but with no inclusion of the supplier. They, as well as other researchers, show how maintenance and inventory policies affect each other and affect the overall inventory system optimality and conclude that, although sequential optimization can give good results, joint optimization adds value to the results.

Vaughan (2005) also addresses inventory policy for service parts when demand for the service parts arises due to scheduled and unscheduled maintenance. A stochastic dynamic programming model is used to characterize an ordering policy that addresses both sources of part demand. In their model, the author assumes a system with *n* identical parts and single-unit demand between preventive maintenance periods. The conclusion is that the cost savings tend to be greatest when there is a small holding cost and a high ordering cost. Furthermore, the optimal policy suggests ordering preventive maintenance units some number of periods prior to the preventive maintenance period.

De Smidt-Destombes *et al.* (2005) develop a model using Markov chains for a system where parts that wear out are managed in two different stages, *i.e.*, degrade and total failure. They assume that the time between the two stages and the repair time are exponentially-distributed. They test their accuracy using discrete-event simulation and find that one of the complications is the strong correlation between the parameters, which makes it difficult to compute the expressions to determine the availability of parts. However, they develop a heuristic to optimize the costs and the availability.

According to Arnold and Köchel (1996), optimal decisions regarding service parts inventories can be found coupling simulation with a search algorithm (*e.g.*, tabu search, simulated annealing, genetic algorithm). Although there is some work done in the inventory optimization area using genetic algorithms (*e.g.*, Köchel and Nieländer, 2002; Pal *et al.*, 2005), limited work exists in joint service parts inventory and maintenance optimization. Carnero (2004) use a genetic algorithm to find near optimal solutions in the development of a methodology for the detection of anomalies in preventive maintenance programs. Marseguerra *et al.* (2004) explore the possibility of using genetic algorithms and Monte Carlo simulation to determine the optimal number of service parts required in storage in a multi-component system. They define two objectives: maximization of revenue in the system and minimization of total spares volume. They conclude that using multiobjective optimization rather than a single objective, where the other objectives are taking in consideration as constraints, yields a more realistic set of results. They also state that the combination of computer simulation and genetic algorithms overcomes the intrinsic limitations of the analytical methods.

We summarize the most closely-related work to the joint optimization of maintenance and service parts inventories, where several models have been developed. However, few of them consider the joint optimization of service part inventories and maintenance.

2.5. Evolutionary Algorithms for Multiobjective Optimization

Many heuristic search algorithms have been developed to solve multiobjective optimization problems including simulated annealing, tabu search, scatter search, ant colony, particle swarm optimization, and evolutionary algorithms (EAs). However, MOEAs have been shown to intelligently balance exploration and exploitation of the solution search space (Deb, 2001). Other advantages of using MOEAs to solve multiobjective problems include:

- EA-based approaches are capable of exploring the search space more thoroughly within a smaller number of solution evaluations than other point-to-point local search procedures (April *et al.*, 2003); and
- EA-based approaches are less dependent on the selection of the starting solutions, and they do not require definition of a neighborhood (April *et al.*, 2003).

Generally, when solving MOOPs, there are three primary goals: (1) fast convergence to the true Pareto frontier solution set in the objective space, (2) close proximity to the true Pareto frontier solution set, and 3) diversity and even dispersion of the nondominated solutions obtained along the true Pareto optimal front. Fast convergence to the set of the true Pareto front and diversity and even dispersion of the set of obtained nondominated solutions for computationally-expensive MOOPs are critical. This is especially the case in real-world problems where finding the optimal or even near-optimal solutions is often computationally-prohibitive.

In recent years, several variations of MOEAs have been developed to handle MOOPs (*e.g.*, Coello *et al.*, 2002; Deb, 2001), including an improved version of the nondominated sorting genetic algorithm (NSGA-II) (Deb *et al.*, 2002). Of these, NSGA-II stands out for its fast nondominated sorting approach, elitism approach, and its overall capability to maintain a better solution spread. Further, it has been reported that NSGA-II outperforms most other MOEAs in terms of convergence to the true Pareto optimal front while maintaining solution diversity. Some studies report that there is no statistically-significant difference between the performance of NSGA-II and other existing MOEAs (Deb *et al.*, 2002; Zitzler *et al.*, 2001). We, therefore, are motivated to use NSGA-II of Deb *et al.* (2002) (with problem-specific modifications) for multiobjective optimization of joint inventory and maintenance policies. In this study, we use NSGA-II but the cost formulations presented here can be used with any existing MOEA that can optimize two or more objectives, including a newer version of strength Pareto EA (SPEA2) (Zitzler *et al.*, 2001), Pareto-archived evolution strategy (PAES) (Knowles and Corne, 1999), rank-density-based multiobjective genetic algorithm (RDGA) (Lu and Yen, 2003), ParEGO (Knowles, 2006), and FastPGA (Eskandari and Geiger, 2006).

2.6. Summary

The problem on which this research focuses spans three research areas in the literature:

- Service parts inventory policy modeling and optimization,
- maintenance policy modeling and optimization, and
- multiobjective optimization.

As discussed here, there is a vast amount of research in these areas. This investigation has found that these areas can be integrated to better represent real-world scenarios. In other words, developing a joint optimization of service parts inventories and maintenance in a supply system is something that needs to be addressed, because these two activities complement each other and have not been examined from that perspective. Moreover, optimizing a multi-member system in which all the members have individual and potentially conflicting objectives is something that has not been largely investigated. Thus, it has been found that there is no research regarding maintenance and service parts inventory simultaneous optimization in a configuration with lateral transshipments, where all the members of the system have separate objectives to be fulfilled.

CHAPTER 3:
JOINT MAINTENANCE AND SERVICE PARTS INVENTORY POLICY MULTIOBJECTIVE OPTIMIZATION FOR A TWO-ECHELON SINGLE-SUPPLIER, *n*-CUSTOMER SUPPLY CHAIN SYSTEM

3.1. Introduction

In this chapter, a joint service parts and maintenance multiobjective optimization model for a simple multi-echelon inventory system is developed. This system configuration allows for tractable analysis while serving as a good starting point for studying larger, more complex real-world service parts inventory systems along with helping us to understand the relationship between service parts inventories and maintenance policies. First, we describe the general configuration and state all the parameters and decision variables included in the model. Next, the long-run expected cost functions for the customers and the supplier are developed. The results of these derivations are used in a single objective and a multiobjective optimization solution approach that are explained later. Finally, this chapter is concluded with experiments and results for a single-customer configuration and a two-customer configuration.

3.2. Description of the Generalized Multi-Echelon Service Part Supply Chain System

Consider a single service part inventory system consisting of one parts supplier that services n customers, as shown in Figure 3.1. The customers are autonomous systems with their own objectives to satisfy, characterized by the typical parameters of reliability models such as failure distributions, part purchase costs, part ordering costs, part replacement costs, equipment downtime costs, service part inventory holding costs, *etc*. The customers want to minimize their long-run expected total maintenance and service parts inventory costs, while the supplier seeks to minimize its long-run expected total part production and inventory cost. The customers must decide the time t_o to order a replacement part and the time t_r to replace a part (before or after it fails) in order to minimize their total maintenance and service parts inventory costs. The supplier must decide the time t_p to start production (manufacture or procurement) of service parts to meet the customers demand for them in order to minimize its total part production and inventory cost. It is important to mention that, if the customer orders a part and the supplier does not have one on-hand for immediate shipment, then the supplier must begin production of the part (or complete production of the part if one is already in process originally intended for stocking) and then arrange an emergency shipment of the part to expedite its delivery to the requesting customer. In most practical cases, it is reasonable to assume that the emergency shipping cost, which is incurred by the supplier, is quite large relative to the inventory holding cost per unit at the supplier.

Now, given this supply system, the long-run expected total cost functions for each independent system in the supply network can be developed. First, the relevant notation, parameters, variables and simplifying assumptions are presented.

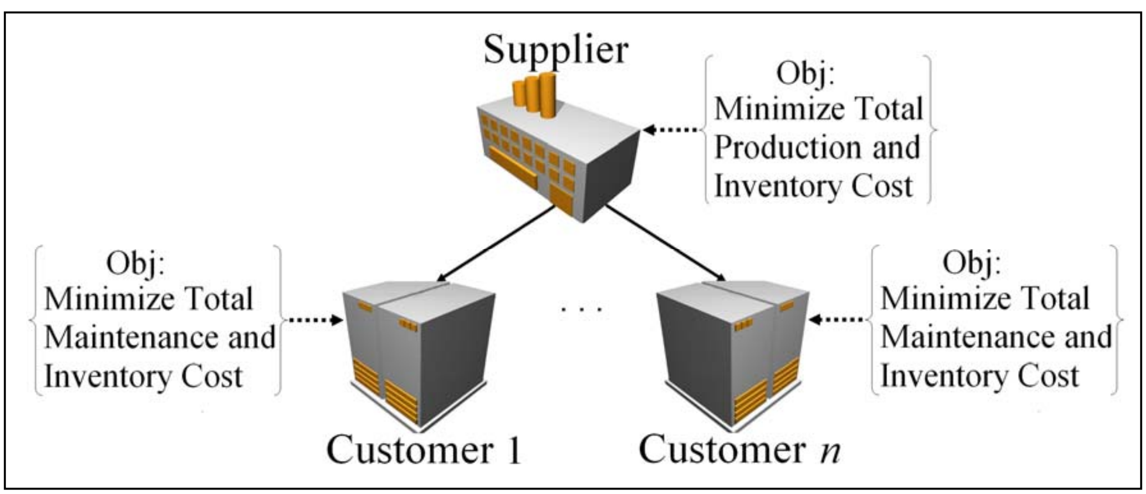

Figure 3.1. The two-echelon, single-supplier, *n*-customer service part inventory system.

Relevant Notation, Parameters and Variables:

$f(t)$: Probability density function of the service part failure distribution as a function of t; $f(t)$ can be any probability density such as Weibull, Gamma, Exponential, Lognormal, *etc.*

Customer-Specific Parameters:

Each customer i in the supply network has these parameters, but each has its own values of these parameters.

L : Lead time for a customer to receive an ordered service part from the supplier; (unit of time, *e.g.*, minutes, hours, days, *etc.*)

c_r : Part replacement cost at a customer; represents all costs incurred when a part is replaced including the cost of the part and the labor cost; (cost per replacement)

c_f : Part failure cost at a customer; represents the costs associated with a failure including repairing any damage that may have occurred at the time of failure and any subsequent damage to the system until the failed part is replaced; (cost per failure)

c_s : Part shortage cost at a customer; (cost per unit time)

c_h^c : Part holding cost at a customer; (cost per unit time)

c_o : Part ordering cost at a customer; represents all costs associated with placing an order and when that order is received, including any clerical/labor costs of processing orders, inspection and return of poor quality units and material handling costs; (cost per order)

Supplier-Specific Parameters:

$g(n)$: Probability mass function of the supplier having a service part in stock as a function of n, the number of customers serviced by the supplier

c_h^s : Part holding cost at the supplier; (cost per unit time)

c_p : Part production cost at the supplier; represents all costs associated with producing a part at the supplier including any production setup cost, or product; (cost per part)

p : Unit production time at the supplier; represents the production lead time if the supplier produces the part in-house or the delivery lead time of the part to the service part supplier from the supplier's supplier if production of the part is outsourced; (unit of time, *e.g.*, minutes, hours, days, *etc.*)

c_e : Emergency shipping cost incurred by the supplier; (cost per unit time)

Decision Variables:

t_o : Scheduled time to order a replacement service part (at the customer)

t_r : Scheduled time to replace a service part (at the customer)

t_p : Scheduled time to produce a replacement service part (at the supplier)

Modeling and Analysis Assumptions:

- Only one service part type is considered within the inventory supply system;
- The probability the supplier will have a part available when requested is a monotonically decreasing function of the number of customers;
- The parts are non-repairable (*i.e.*, replacement);

- Only one part is ordered at a time and a maximum of one service part is held in inventory by a customer, *i.e.*, base stock model;
- If a customer has service part inventory on-hand, then a new order cannot be placed by that customer;
- The time between successive part replacements at a customer is considered a cycle for the customer;
- The time between successive part production runs at the supplier is considered a cycle for the supplier;
- Lead time L to receive a replacement part from the supplier after an order is placed is positive and fixed;
- Actual part replacement once the part is received after L is assumed to be instantaneous and perfect;
- If a part is needed from the supplier, the time at which a part is replaced cannot occur before the part arrives from the supplier, *i.e.*, $t_o + L + p \le t_r$;
- No lateral transshipments of service parts occur between customers;
- All activities required to expedite a customer order are assumed by the supplier; and
- All the costs are positive.

3.3. Modeling and Solving the Customer's Problem

The analysis begins by constructing the long-run expected total cost function for a customer in the service part supply network. Assuming a cycle at each customer is the length of time between successive part replacements, we use renewal theory and the renewal-reward process with arbitrary inter-renewal times to compute the long-run expected cost per unit time. In other words, suppose we have a sequence of independent and identically distributed random cycle lengths $\{X_n: n \geq 1\}$. If random variable Y_n is the expected total cost incurred during interval n and $\{Y_n: n \geq 1\}$, then the long-run expected cost per unit time at the customer is expressed as $E[TC_c] = \lim_{t\to\infty} \frac{1}{t} Y_t = E[X]/E[Y]$. Hence, $E[TC_c]$ is equal to the expected costs divided by the expected length of the cycle, and the aim is to minimize $E[TC_c]$. For details of renewal theory and the renewal-reward process, the reader is referred to Medhi (1982) and Feldman and Valdez-Flores (1995).

The long-run expected total cost associated with each customer includes expected costs due to ordering, maintenance, inventory holding, and inventory shortages as well as the expected cycle length, that is,

$$E[TC_c] = \frac{\begin{array}{c}\text{Ordering Cost} + E[\text{Maintenance Cost}] \\ + \, E[\text{Inventory Holding Cost}] \, + \, E[\text{Inventory Shortage Cost}]\end{array}}{E[\text{Cycle Length}]}. \tag{3.1}$$

The ordering cost c_o is incurred in every cycle since only one part is ordered at a time and a maximum of one service part is held in inventory by a customer. The expected maintenance cost is simply the sum of the unit replacement cost c_r and the expected cost due to a failure before the part is replaced. Thus,

$$E[\text{Maintenance Cost}] = c_r + c_f \int_0^{t_r} f(t)\,dt. \tag{3.2}$$

If a replacement part is ordered and arrives before the existing part fails, then inventory holding costs are incurred from the time the replacement part arrives until it is used for preventive or corrective maintenance. In addition, the holding cost includes the probability of the part not failing. So, if the part is received, it is held in inventory. The holding cost also reflects whether the supplier has or does not have the part in stock ready for immediate shipment. This is given by $g(n)$ if the part is in stock and 1-$g(n)$ if the part is not in stock. If the supplier does not have a part in stock, then the customer has to account for p additional units of time for the supplier to produce the part. The expected holding cost is

$$E\left[\text{Holding Cost}\right]=c_h^c\begin{bmatrix} g\left(n\right)\left(\int_{t_o+L}^{t_r}\left(t-t_o-L\right)f\left(t\right)dt+(t_r-t_o-L)\int_{t_r}^{\infty}f\left(t\right)dt\right)+ \\ \left(1-g\left(n\right)\right)\left(\int_{t_o+L+p}^{t_r}\left(t-t_o-L-p\right)f\left(t\right)dt+(t_r-t_o-L-p)\int_{t_r}^{\infty}f\left(t\right)dt\right)\end{bmatrix}. \quad (3.3)$$

On the other hand, if a replacement part is ordered after an existing part fails, then shortage costs are incurred during the lead time associated with processing the customer's order (*i.e.*, the time from part failure to the time the customer receives the replacement part). Shortage costs can also be incurred if an existing part fails after the order for a replacement part has been received by the supplier, but before it arrives at the customer's location from the supplier. Finally, shortage costs can continue to accrue for the additional time associated with production (or procurement) if the supplier does not have a service part on hand to ship to a requesting customer after that customer's order is received. Under these conditions, the customer's expected shortage cost can be expressed as

$$E\left[\text{Shortage Cost}\right]=c_s\begin{bmatrix} g\left(n\right)\left(\int_{0}^{t_o}L\,f\left(t\right)dt+\int_{t_o}^{t_o+L}\left(t_o+L-t\right)f\left(t\right)dt\right) \\ \left(1-g\left(n\right)\right)\left(\int_{0}^{t_o}(L+p)f\left(t\right)dt+\int_{t_o}^{t_o+L+p}\left(t_o+L+p-t\right)f\left(t\right)dt\right)\end{bmatrix}. \quad (3.4)$$

The expected cycle length for the customer is expressed as

$$E[\text{Cycle Length}] = g(n)\left[\int_0^{t_o}(t+L)\,f(t)\,dt + \int_{t_o}^{t_o+L}(t_o+L)\,f(t)\,dt + \int_{t_o+L}^{t_r} tf(t)\,dt + \int_{t_r}^{\infty} t_r\,f(t)\,dt\right]$$
$$+\left(1-g(n)\right)\begin{bmatrix}\displaystyle\int_0^{t_o}(t+L+p)\,f(t)\,dt + \int_{t_o}^{t_o+L+p}(t_o+L+p)\,f(t)\,dt \\ \displaystyle + \int_{t_o+L+p}^{t_r} tf(t)\,dt + \int_{t_r}^{\infty} t_r\,f(t)\,dt\end{bmatrix}. \tag{3.5}$$

In this study, we assume that $g(n)$ is a nondecreasing function of the number of customers n that the supplier currently services. In other words, as the number of customers increases, the probability of the supplier not having a part available when requested by a customer in the supply network increases. Substituting Eqs. 3.2–3.5 into Eq. 3.1, the customer's long-run expected total cost per unit time is

$$E[TC_c] = \frac{\begin{bmatrix} \displaystyle c_o + c_r + c_f\int_0^{t_r} f(t)\,dt + c_h^c g(n)\left(\int_{t_o+L}^{t_r}(t-t_o-L)\,f(t)\,dt + (t_r-t_o-L)\int_{t_r}^{\infty} f(t)\,dt\right) + \\ \displaystyle c_h^c\left(1-g(n)\right)\left(\int_{t_o+L+p}^{t_r}(t-t_o-L-p)\,f(t)\,dt + (t_r-t_o-L-p)\int_{t_r}^{\infty} f(t)\,dt\right) + \\ \displaystyle c_s g(n)\left(\int_0^{t_o} Lf(t)\,dt + \int_{t_o}^{t_o+L}(t_o+L-t)\,f(t)\,dt\right) + \\ \displaystyle c_s\left(1-g(n)\right)\left(\int_0^{t_o}(L+p)\,f(t)\,dt + \int_{t_o}^{t_o+L+p}(t_o+L+p-t)\,f(t)\,dt\right)\end{bmatrix}}{\begin{bmatrix} \displaystyle g(n)\left(\int_0^{t_o}(t+L)\,f(t)\,dt + \int_{t_o}^{t_o+L}(t_o+L)\,f(t)\,dt + \int_{t_o+L}^{t_r} tf(t)\,dt + \int_{t_r}^{\infty} t_r\,f(t)\,dt\right) + \\ \displaystyle \left(1-g(n)\right)\begin{pmatrix}\displaystyle\int_0^{t_o}(t+L+p)\,f(t)\,dt + \int_{t_o}^{t_o+L+p}(t_o+L+p)\,f(t)\,dt + \\ \displaystyle\int_{t_o+L+p}^{t_r} tf(t)\,dt + \int_{t_r}^{\infty} t_r\,f(t)\,dt\end{pmatrix}\end{bmatrix}}. \tag{3.6}$$

Derivation of the Customer's Inventory and Maintenance Policies

The exponential distribution is perhaps the simplest life distribution model. In reliability science, this distribution forms the basis for many comparisons and is widely used in reliability specifications. The density for the exponential probability distribution is

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & \text{otherwise,} \end{cases} \tag{3.7}$$

and the cumulative distribution is

$$\begin{aligned} F(t) &= \int_0^t \lambda e^{-\lambda t} dt \\ &= 1 - e^{-\lambda t}. \end{aligned} \tag{3.8}$$

Using Eq. 3.6 and Eq. 3.7, the reliability function $R(t)$ can be expressed as

$$\begin{aligned} R(t) &= \frac{f(t)}{1 - F(t)} = \frac{\lambda e^{-\lambda t}}{1 - \left[1 - e^{-\lambda t}\right]} \\ R(t) &= \lambda. \end{aligned} \tag{3.9}$$

The parameter λ is the constant failure rate measured in the number of failures per unit time and is equal to μ^{-1}, *i.e.*, $\lambda = 1/\mu$, where μ is the mean time between part failures. Due to its constant failure rate property, the exponential distribution is a reasonable model for the long flat "intrinsic failure" portion of the well-known reliability bathtub curve. Since most components and systems spend most of their lifetimes in this portion of the bathtub curve, this perhaps explains the frequent use of the exponential distribution (when an early failure during infant mortality or wear out is not a concern). In the case of service parts, it is reasonable to assume exponential failures when the parts are replaced and not repaired, which is the case in this research. When parts are replaced, the original failure rate of the part can be assumed to remain the same. In real case scenarios, data would have to be collected to determine the actual failure distribution of the part.

Assuming that the random part failures follow a stationary exponential distribution, we now find the time at which to order a service part (t_o) and the time at which to replace a part (t_r) for the customer. Substituting Eq. 3.7 into Eq. 3.6 and simplifying, the long-run expected total cost per unit time for the customer assuming an exponential part failure distribution is

$$E\left[TC_c\right]=\frac{\left[\begin{array}{l} c_r+c_f+c_f\,e^{-\lambda t_r}+c_o+ \\ c_h^c\left(\dfrac{e^{-\lambda\,(2L+p+2t_o+t_r)}\left(e^{\lambda\,(L+p+t_o+t_r)}g(n)-e^{\lambda\,(2L+p+2t_o)}-e^{\lambda\,(L+t_o+t_r)}\left(g(n)-1\right)\right)}{\lambda}\right)+ \\ c_s\left(L+p-g(n)\,p-\dfrac{e^{-\lambda\,(L+p+t_o)}\left(e^{\lambda\,(L+p)}+g(n)\;-\;e^{\lambda\,p}\left(g(n)-1\right)\right)}{\lambda}\right) \end{array}\right]}{\left(L+p-g(n)\,p-\dfrac{e^{-\lambda t_o}+e^{-\lambda t_r}+e^{-\lambda(L+p+t_o)}\left(g(n)-1\right)-e^{-\lambda(L+t_o)}\left(g(n)-1\right)}{\lambda}\right)}. \tag{3.10}$$

The optimal values of t_o and t_r are found by setting the partial derivatives $\partial E\left[TC_c\right]/\partial t_o$ and $\partial E\left[TC_c\right]/\partial t_r$ equal to zero and solving the resulting equations simultaneously. It follows that the optimal values are

$$t_o^*=\frac{\ln\left(-\dfrac{e^{-\lambda\,(L+p)}\left(\left(1+g(n)\left(e^{\lambda p}-1\right)\right)\left(c_f\lambda-c_s\right)-e^{\lambda\,(L+p)}\left(c_h^c+c_f\lambda-c_s\right)\right)}{\lambda\left(c_o+c_r\right)+c_h^c\left(1+\lambda\left(L+p-g(n)\,p\right)\right)+\lambda\left(L+p-g(n)\,p\right)\left(c_f\lambda-c_s\right)}\right)}{\lambda} \quad \text{and} \tag{3.11}$$

$$t_r^*=\frac{\ln\left(\dfrac{e^{\lambda\,(L+p)}\left(c_h^c+c_f\lambda-c_s\right)+\left(\left(1+g(n)\left(e^{\lambda p}-1\right)\right)\left(-c_f\lambda+c_s\right)\right)}{\begin{array}{l} c_h^c\left(g(n)-e^{\lambda p}g(n)-1\right)\left(\lambda p\left(g(n)-1\right)-L\lambda-1\right)+\left(e^{\lambda\,(L+p)}+g(n)-e^{\lambda p}g(n)-1\right) \\ \left(\lambda\left(c_o+c_r+c_f\right)-c_s\right)\end{array}}\right)}{\lambda}. \tag{3.12}$$

In order to determine the convexity of the cost functions over the two variables t_o and t_r, we compute the second derivatives, which corresponds to the determinant of the Hessian matrix

$$H\left(E\left[TC_c\right]\right)=\begin{vmatrix} \dfrac{\partial^2 E\left[TC_c\right]}{\partial t_o{}^2} & \dfrac{\partial^2 E\left[TC_c\right]}{\partial t_o \partial t_r} \\ \dfrac{\partial^2 E\left[TC_c\right]}{\partial t_r \partial t_o} & \dfrac{\partial^2 E\left[TC_c\right]}{\partial t_r{}^2} \end{vmatrix}.$$

We need to find a set of points where $H(E[TC_c])$ is positive, which indicates that the multivariate function is convex over that area and a global minimum can be found as a function of both t_o and t_r. We could not show that there is such a region, which means that the function is not strictly convex over the two decision variables. Given the complexity of the determinant, the calculations are not shown in this document. However, following the analysis performed by Armstrong and Arkins (1996) who use similar models, we can say that the long-run expected total cost per unit time function is pseudoconvex in t_r. This means, for a given t_o, the optimal t_r can be either $(t_o + L + p)$, ∞, or t_r^*. In the case that t_r^* does not exist, the partial derivative $\partial E\left[TC_c\right]/\partial t_r$ determines when to replace the part. If it is negative, then the part should be replaced at failure. If it is positive, then the part should be replaced at the instant the part is received from the supplier at $(t_o + L + p)$. Similarly, the long-run expected total cost per unit time function is pseudoconvex in t_o. So, for a given t_r, the optimal t_o can be either 0, $(t_r – L – p)$, or t_o^*. If t_o^* does not exist, then the time to order is determined by the sign of the partial derivative $\partial E\left[TC_c\right]/\partial t_o$. If it is positive, the customer should place an order at time $(t_r – L – p)$. If it is negative, then the customer should place an order at $t = 0$.

3.4. Modeling and Solving The Supplier's Problem

The supplier's long-run expected cost per unit time structure consists of production, inventory holding, and emergency production/shipping costs divided by the expected cycle length. In other words,

$$E\left[TC_s\right]=\frac{\text{Production Cost}+E[\text{Inventory Holding Cost}]\ +\ E[\text{Emergency Shipping Cost}]}{E[\text{Cycle Length}]}. \quad (3.13)$$

Production cost is incurred in every cycle when a part is produced. Holding cost at the supplier is incurred any time a part is produced but an order from the customer has not, yet, been placed. Emergency shipping cost is incurred anytime a shipment is required when a part is not on-hand and ready to be shipped when ordered by a customer. This cost is essentially considered a shortage cost for the supplier. Holding and emergency costs are a function of $g(n)$ and $1-g(n)$, respectively. The supplier's combined production, holding and emergency costs E[PHE Cost] is expressed as

$$\begin{aligned} E[\text{PHE Cost}]=c_p+c_h^s g(n)&\left(\int_{t_p+p}^{t_o}\left(t_o-t\right)f(t)\,dt+\left(t_o-p-t_p\right)\int_{t_o}^{\infty}f(t)\,dt\right)+ \\ c_e\left(1-g(n)\right)&\left(\int_{t_o}^{t_o+p}pf(t)\,dt\right). \end{aligned} \quad (3.14)$$

In this problem, the supplier seeks to minimize the total long-run expected cost per unit time. Therefore, the total long-run expected cost rate is equal to the total cost divided by the expected cycle length. The cycle length for the supplier is

$$E\left[\text{Cycle Length}\right]=g(n)\left(\int_0^{t_o}(t)\,f(t)\,dt+\int_{t_o}^{\infty}(t_o)\,f(t)\,dt\right)+$$
$$\left(1-g(n)\right)\left(\int_0^{t_p}(t+p)\,f(t)\,dt+\int_{t_p}^{t_o}\left(t_p+p\right)\,f(t)\,dt+\int_{t_o}^{\infty}(t_o+p)\,f(t)\,dt\right). \tag{3.15}$$

Substituting Eqs.3.14 and 3.15 into Eq. 3.13 the supplier's long-run expected total cost per unit time is

$$E[TC_s]=\frac{\begin{bmatrix} c_p+c_h^s g(n)\left(\displaystyle\int_{t_p+p}^{t_o}(t_o-t)\,f(t)\,dt+\left(t_o-p-t_p\right)\int_{t_o}^{\infty}f(t)\,dt\right)+ \\ c_e\left(1-g(n)\right)\left(\displaystyle\int_{t_o}^{t_o+p}pf(t)\,dt\right)\end{bmatrix}}{\begin{bmatrix} g(n)\left(\displaystyle\int_0^{t_o}tf(t)\,dt+\int_{t_o}^{\infty}t_o\,f(t)\,dt\right)+ \\ \left(1-g(n)\right)\left(\displaystyle\int_0^{t_p}(t+p)\,f(t)\,dt+\int_{t_p}^{t_o}\left(t_p+p\right)\,f(t)\,dt+\int_{t_o}^{\infty}(t_o+p)\,f(t)\,dt\right)\end{bmatrix}}. \tag{3.16}$$

<u>Derivation of the Supplier Inventory and Production Policies</u>

The expected total cost per unit time for the supplier (Eq 3.16) assuming an exponential part failure distribution is

$$E[TC_s]=\frac{\begin{bmatrix} e^{\lambda\,(t_o+t_p)}\lambda\left(c_p-c_e e^{-\lambda\,(t_o+p)}p\left(e^{\lambda p}-1\right)\left(g(n)-1\right)\right)+ \\ \left(\dfrac{1}{\lambda}\right)\left(c_h^s\,g(n)e^{-\lambda\,(t_o+t_p+p)}\left(e^{\lambda t_o}\left(-\lambda\left(p-t_o+t_p\right)-1\right)+e^{\lambda\,(t_p+p)}\left(1-\lambda\left(p-t_o+t_p\right)\right)\right)\right)\end{bmatrix}}{e^{\lambda\,t_o}\left(g(n)-1\right)+e^{\lambda\,t_p}\left(\lambda\left(t_o-t_p\right)+g(n)\left(t_p\lambda-t_o\lambda-1\right)\right)+e^{\lambda\,(t_o+t_p)}\left(1+p\left(\lambda-g(n)\right)\lambda\right)}. \tag{3.17}$$

Similar to the customer's problem, the optimal values of t_o and t_p can be found by setting the partial derivatives $\partial E\left[TC_s\right]/\partial t_o$ and $\partial E\left[TC_s\right]/\partial t_p$ equal to zero and solving the resulting equations simultaneously, yielding

$$t_o^* = \frac{\ln\left(\dfrac{e^{-\lambda p}\left(c_h^s\, g(n)\, e^{2\lambda p} + c_e\, p\, \lambda\left(e^{\lambda p}-1\right)^2\left(g(n)-1\right) - c_h^s\left(g(n)-1\right)\left(1+p\lambda\right)\right)}{c_h^s + c_p\, \lambda\, \left(e^{\lambda p}-1\right) + c_h^s\, p\, \lambda}\right)}{\lambda} \text{ and} \tag{3.18}$$

$$t_p^* = \frac{\ln\left(\dfrac{g(n)\left(c_h^s\, g(n)\, e^{2\lambda p} + c_e\, p\, \lambda\left(e^{\lambda p}-1\right)^2\left(g(n)-1\right) - c_h^s\left(g(n)-1\right)\left(1+p\lambda\right)\right)}{\begin{array}{l} e^{\lambda p}\, \lambda\, \left(g(n)-1\right)\left(c_p - p\, c_e\right)\left(1+p\,\lambda\right) + \\ e^{2\lambda p}\left(c_h^s\, \left(g(n)+g(n)\, p\, \lambda\right) + \lambda\left(c_p\, g(n) + c_e\, p\left(g(n)-1\right)\left(1+p\,\lambda\right)\right)\right) \end{array}}\right)}{\lambda}, \tag{3.19}$$

respectively.

Similar to the customer's problem, no conclusion is made about the convexity of the function in terms of t_o and t_p based on the second derivative test. The long-run expected total cost per unit time function for the supplier is pseudoconvex in t_o. Therefore, for a given t_p, the optimal t_o can be either (p + t_p), ∞ or t_o^*. In the case that t_o^* does not exist, the partial derivative $\partial E\left[TC_s\right]/\partial t_o$ determines the best time at which to receive an order for a replacement service part from a customer. If this derivative is negative, then the best time at which the supplier should receive a customer's order is at time zero. If it is positive, the best time to receive a customer's order is at $t_p + p$. The long-run expected total cost function is pseudoconvex in t_p. So, for a given t_o, which is mostly the case, the optimal t_p can be either 0, ($t_o - p$) or t_p^*. Once more, if t_p^* does not exist, then the time to produce a part is determined by the sign of the partial derivative $\partial E\left[TC_s\right]/\partial t_p$. If it is positive, the supplier should produce a part at ($t_o - p$). On the other hand, if it is negative, the supplier has to produce a part at t = 0. The supplier could potentially have multiple t_p^*'s, one for each of the n customers the supplier services. Then, the problem for the supplier effectively becomes a production scheduling problem for the different customer orders. Integrating this scheduling problem is worthy of and left for further study.

In this section, the joint inventory-maintenance optimization problem within the two-echelon supply network is decoupled into the customer-level problem and the supplier-level problem. Each problem is solved separately and independently from the other. A decoupled optimization is performed using an optimization algorithm with the derived individual total cost functions. Next, a computational study is conducted to determine the system-wide cost of the supply chain. First, each cost function is optimized separately. Then, using a multiobjective optimization approach, the cost functions are optimized simultaneously. The results of both approaches are compared. Before presenting the computational study, in Section 3.5, an overview of the optimization algorithms that are used is given.

3.5. Decoupled and Multiobjective Optimization Solution Approaches

3.5.1. Real-Coded Genetic Algorithm for Single Objective Optimization

$E[TC_s]$ and $E[TC_c]$ are the objective functions that are utilized by the chosen optimization algorithms for decoupled and simultaneous multiobjective optimization. In this research, for the decoupled optimization, we use a real-coded genetic algorithm (RCGA) to generate the inventory ordering and maintenance policies (*i.e.*, t_o, t_r and t_p). As GAs are popular optimization procedures and common in many different applications, we forgo describing this popular optimization algorithm and refer the reader to the works of Goldberg (1989) and Pal *et al.* (2005) for details of the RCGA.

3.5.2. Real-Coded NSGA-II for Multiobjective Optimization

For multiobjective optimization, we employ the real-coded NSGA-II (Deb *et al.*, 2002) to generate the joint policies. However, the expected total cost per unit time formulations are quite suitable for using other EA and MOEA nondomination-based optimization approaches. NSGA-II is an improved, elitism version of NSGA by Srinivas and Deb (1994), where the fitness of a solution is obtained by a Pareto ranking procedure. NSGA-II starts with an initial, random population of solutions P_0 of size N. This initial population is then sorted based on nondomination, which means that none of the solutions is better than the others with respect to all objectives. At this point, each solution is assigned a fitness value equal to its domination level, where 1 corresponds to the nondomination level, 2 is the next best (dominated) level, and so on. The first level contains solutions that dominate solutions of all other levels. The nondominating sorting algorithm uses this fitness value to rank the solutions and assign them to the different fronts. Each solution belongs to different fronts based on its domination level – Front F_1 = Level 1, Front F_2 = Level 2, *etc*. Then, using the evolutionary algorithm operators of binary tournament selection, simulated binary crossover operator, and polynomial mutation, an offspring population Q_0 of size N is created.

Beginning with the first generation i = 1 of the algorithm, the procedure for the ith population is different than that for the initial population. First, the offspring population Q_i is combined with the parent population to create a combined population $R_i = P_i \cup Q_i$ of size $2N$. Then, this new population R_i is sorted according to nondomination. This allows the parent solutions to be compared with the child population, thereby ensuring elitism. This sorting classifies the population into several fronts F_1, F_2, F_3, and so on. All solutions belonging to the best nondominating frontier set (F_1) are emphasized more than any other solution in the combined population. If the size of F_1 is smaller than N, then all members of F_1 are chosen for the new population P_{i+1}. The remaining members of the population are chosen from the subsequent nondominated fronts in the order of their ranking. In other words, solutions from front F_2 are chosen next, followed by solutions from front F_3, and so on. This process is repeated until no more fronts can be accommodated and the size of population P_{i+1} reaches N.

In general, the count of solutions in all fronts from F_2 to F_l would be larger than the population size. To choose exactly N population members, the population of the last front is sorted with the crowding-distance operator. This operator facilitates the selection of the best solutions to fill the population slots by using the average distance of two solutions along each of the objectives. The binary tournament selection uses the crowded-distance to choose one of the two solutions. Between two solutions on different fronts in a tournament, solutions with lower rank are preferred. Otherwise, if both the solutions belong to the same front then the solution that is located in a region with fewer number of other solutions (*i.e.*, with a larger crowded distance) is preferred. As a result, solutions from less dense regions in the search space are given importance in deciding which solutions to choose from R_i to construct population P_{i+1}. This operator helps to fill population P_{i+1} up to size N. Since the overall population size of R_t is $2N$, not all fronts may be accommodated in N slots available in the new population. All members of fronts that are not selected for the next population using the crowded-distance operator are simply deleted. Complete details of NSGA-II can be found in Deb (2001) and Deb *et al.* (2002).

3.6. Computational Study

This section shows the performance of the decoupled optimization and multiobjective optimization approaches for joint inventory and maintenance policy optimization. Under the decoupled approach, each member of the service parts supply chain seeks to minimize its own long-run expected total cost per unit time without consideration or collaboration with the other members of the supply network. Customers in the network seek to minimize their long-run expected total maintenance and inventory costs, and the supplier seeks to minimize its long-run expected total production and inventory cost per unit time. Based on their individual costs, the system-wide cost per unit time for the supply chain network can be determined. Under the multiobjective modeling approach, the optimal values of t_o, t_r and t_p that minimize the total system-wide cost per unit time are identified. Two supply chain configurations are examined. The first configuration considers only a single parts supplier and a single customer (n = 1). The second case considers a single parts supplier and two customers (n = 2).

3.6.1. Single-Supplier, Single-Customer, Multi-Echelon Service Parts Inventory System

Table 3.1 summarizes the experimental design of the problem instances for the single-supplier, single-customer supply chain configuration. The values and ranges of the parameters are chosen somewhat arbitrarily. For the single service part, a failure distribution of exponential form with a rate parameter $\lambda = 0.01$, or a mean time between part failures $\mu = 100$, is used.

Table 3.1. Ranges of the problem instance parameters for a two-echelon, one-supplier, one-customer service parts inventory supply chain system.

	Supplier	Customer
Unit Ordering Cost, c_o	-	10
Unit Holding Cost, c_h^s, c_h^c	[300, 600]	[300, 600]
Unit Replacement Cost, c_r	-	30
Unit Failure Cost, c_f	-	28
Unit Shortage Cost, c_s	-	[300, 600]
Mean Time Between Failure, μ	-	100
Unit Emergency Shipping Cost, c_e	[300, 600]	-
Unit Production Cost, c_p	50	-
Unit Production Time, p	4	-
Order Delivery Lead Time, L	2	-

3.6.1.1. Decoupled Optimization Approach

The search control parameters for RCGA are summarized in Table 3.2, which lists the values of the parameters. Via a small pilot parametric study, a population size of 200 is chosen because it shows reasonable convergence behavior. The number of generations, G, is set to 10,000. The crossover rate p_c and mutation rate p_m for this study are set to 1.00 and 0.01, respectively. With these values and the values described in Table 3.1, we generate the part replacement and part ordering policies for the customers at the different parameter levels (resulting in 17 problem instances) and the optimal part production schedule at the supplier (Table 3.3). Using the RCGA, we determine the lowest possible cost that the algorithm can find given the parameters established for each of the 17 problem instances. In Table 3.4, the total cost per unit time values and corresponding t_o, t_r and t_p values are reported.

Table 3.2. Search control parameters for the RCGA.

Parameter	Value
Population Size, P	200
Number of Generations, G	10,000
Crossover Rate, p_c	1.000
Mutation Rate, p_m	0.01

Table 3.3. Specific problem instance cost parameters for a two-echelon, one-supplier two-customer service parts inventory supply chain system.

Prob	Supplier			Customer				
Inst	c_p	c_h^s	c_e	c_r	c_f	c_o	c_h^c	c_s
1	50	300	300	30	28	10	300	300
2	50	600	300	30	28	10	300	300
3	50	300	600	30	28	10	300	300
4	50	600	600	30	28	10	300	300
5	50	300	300	30	28	10	600	300
6	50	600	300	30	28	10	600	300
7	50	300	600	30	28	10	600	300
8	50	600	600	30	28	10	600	300
9	50	300	300	30	28	10	300	600
10	50	600	300	30	28	10	300	600
11	50	300	600	30	28	10	300	600
12	50	600	600	30	28	10	300	600
13	50	300	300	30	28	10	600	600
14	50	600	300	30	28	10	600	600
15	50	300	600	30	28	10	600	600
16	50	600	600	30	28	10	600	600
17	50	450	450	30	28	10	450	450

Table 3.4. Service part ordering, part replacement and part production times when each supply chain member's cost objective is optimized separately under the decoupled approach.

Prob Inst	t_o	t_r	t_p	Supplier Cost/ Time	Customer Cost/ Time	Long-Run Total Cost/ Time
1	4240.17	4530.69	4405.28	214.07	174.34	388.40
2	4240.17	4530.69	4405.28	215.44	174.34	389.77
3	4240.17	4530.69	4405.28	215.44	174.34	389.77
4	4240.17	4530.69	4405.28	215.44	174.34	389.77
5	4240.17	4530.69	4548.60	215.44	174.34	389.77
6	4240.17	4530.69	4548.60	215.44	174.34	389.77
7	4240.17	4530.69	4548.60	215.44	174.34	389.77
8	4240.17	4530.69	4548.60	215.44	174.34	389.77
9	4240.17	4530.69	4405.28	410.89	174.34	585.23
10	4240.17	4530.69	4405.28	410.89	174.34	585.23
11	4240.17	4530.69	4405.28	410.89	174.34	585.23
12	4240.17	4530.69	4405.28	410.89	174.34	585.23
13	4240.17	4530.69	4548.60	410.89	174.34	585.23
14	4240.17	4530.69	4548.60	410.89	174.34	585.23
15	4240.17	4530.69	4548.60	410.89	174.34	585.23
16	4240.17	4530.69	4548.60	410.89	174.34	585.23
17	4240.17	4530.69	4548.60	313.16	174.34	487.50

3.6.1.2. Joint Optimization Approach

The search control parameters for the real-coded NSGA-II, also chosen via on a small pilot study, are summarized in Table 3.5. In the pilot study, the population size is set to 25, 50, 100, 200, 500 and 1000, and it is observed that using a population size of 200 yields good results without suffering the problems of getting trapped at local optima or compromising the spread of the set of Pareto optima. The number of generations, G, is set to 10,000. Crossover probability, p_c, is varied 0.50-1.00 keeping all the other parameters constant. Low probability leads to poorly converged solutions, and the diversity of the solutions is poor. However, beyond p_c = 0.85, solutions are stable with changes in p_c. Mutation probability p_m is varied between 0.00 and 0.30. Low values and high values of the mutation probability cause loss of solution set diversity. It is found that when p_m ranges from 0.15 to 0.25 diversity is preserved as well as convergence. The crossover rate p_c and mutation rate p_m in this study are set to 1.00 and 0.167, respectively.

Table 3.5. Search control parameters for NSGA-II.

Parameter	Value
Population Size, P	200
Number of Generations, G	10,000
Crossover Rate, p_c	1.000
Mutation Rate, p_m	0.167
Distribution Index for Crossover η_c	10
Distribution Index for Mutation η_m	25

The settings of the crossover distribution index η_c and the mutation distribution index η_m are also varied. The crossover distribution index η_c is a positive real number in the range of [5.0, 100.0], in general, and controls the spread of offspring solutions (Deb *et al.* 2002). It has been observed that larger values of η_c give a higher probability for creating solutions "close" to the parents, and smaller values of η_c allow distant solutions to be selected as offspring. It has been observed that in case of highly nonlinear responses, smaller values of η_c yield better results (avoid suboptimal distribution of nondominated points). The mutation distribution index η_m controls the spread of mutated solutions and is generally kept in the range [5.0, 100.0]. For highly nonlinear responses, smaller values of η_m should be used to increase the spread of solutions.

General Behavior of the Set of Pareto Optima

The behavior of the set of Pareto optima follows what would be expected. Figure 3.3 and Figure 3.4 show the efficient frontier for Customer total cost per unit time vs. the Supplier total cost per unit time for two instances of the problem (Prob Inst 10 and Prob Inst 16). There is a reasonable level of diversity in the solutions since there is somewhat uniform spacing of the solutions along the Pareto front with some gaps.

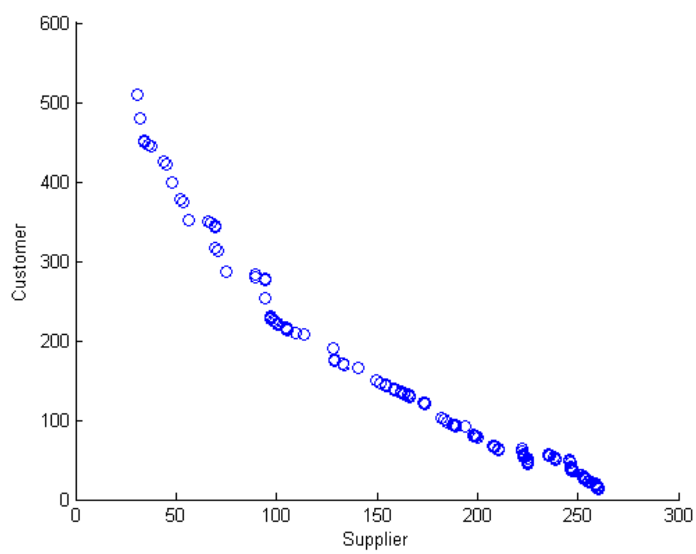

Figure 3.2. Customer long-run expected total cost per unit time versus Supplier long-run expected total cost per unit time for Prob Inst 10.

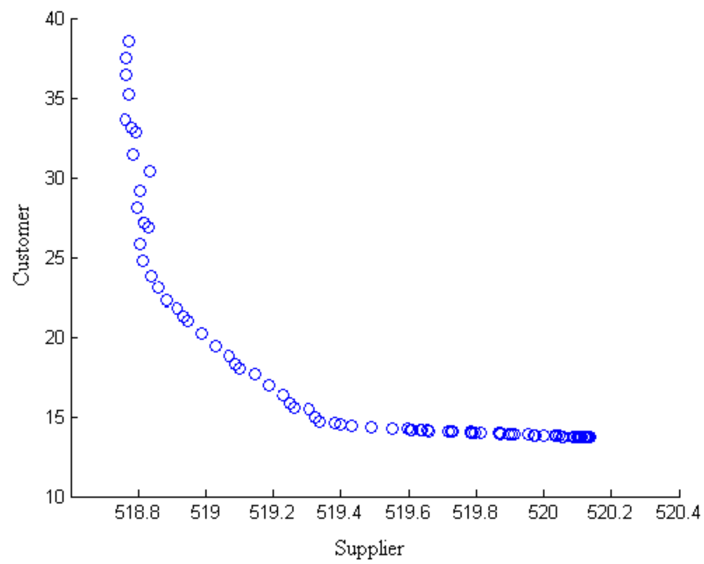

Figure 3.3. Customer long-run expected total cost per unit time versus Supplier long-run expected total cost per unit time for Prob Inst 16.

Using NSGA-II, we determine a set of nondominated solutions for each of the 17 problem instances. Of the set of Pareto optimal solutions for a problem instance, one solution has the lowest total system-wide cost per unit time. The t_o, t_r and t_p values that generated the lowest total cost per unit time for each problem instance are also reported. In order to assess the reasonableness of these results, we use the same problem parameters used for the decoupled optimization (Table 3.1). The results for the joint optimization approach are summarized in Table 3.6. The average improvement over the decoupled optimization over the 17 problem instances is 43.52% with a standard deviation of 20.33%, ranging from 8.79% (Prob Inst 16) to 70.48% (Prob Inst 9). It is clear from Table 3.6 that there is economic benefit to simultaneously optimizing objectives of the individual members of the supply chain network. Pareto optimization permits not only the identification of the set of nondominated solutions where improvement in each solution does not cause any degradation in the quality of the other solutions, but it also permits the identification of the best-compromised solution in terms of the lowest system-wide cost per unit time.

Table 3.6. Service part ordering, part replacement and part production times when each supply chain member's cost per unit time objective is optimized simultaneously under the joint approach.

Prob Inst	t_o	t_r	t_p	Supplier Cost/Time	Customer Cost/Time	Long-Run Total Cost/Time	% Improve Over Decoupled
1	4.09	10.10	0.09	17.94	152.08	170.02	56.23%
2	4513.38	4896.31	3893.24	260.32	7.18	267.50	31.37%
3	4.00	10.00	0.00	13.72	154.69	168.41	56.79%
4	4409.15	4830.20	4182.81	260.32	7.18	267.50	31.37%
5	5.37	12.86	1.11	136.60	179.86	316.46	18.81%
6	4.12	10.23	0.10	26.38	294.98	321.36	17.55%
7	5.24	11.67	1.18	127.78	145.98	273.76	29.76%
8	4.11	10.17	0.09	26.24	291.39	317.63	18.51%
9	4.01	10.01	0.00	13.00	159.78	172.78	70.48%
10	4069.32	4176.39	3967.36	260.32	13.70	274.01	53.18%
11	4.01	10.01	0.01	14.09	159.82	173.91	70.28%
12	4178.29	4592.54	4100.77	260.32	13.70	274.01	53.18%
13	4.28	10.34	0.27	44.22	156.03	200.25	65.78%
14	4.47	10.47	0.38	63.09	280.52	343.60	41.29%
15	4.28	10.34	0.27	44.22	156.03	200.25	65.78%
16	387.68	393.71	382.40	519.66	14.12	533.78	8.79%
17	4.01	10.04	0.01	14.09	226.39	240.47	50.67%

3.6.2. Single-Supplier, *n*-Customer, Multi-Echelon Service Parts Inventory System

The performance of the decoupled optimization approach and the performance multiobjective modeling approach for the single-supplier, *n*-customer multi-echelon service parts inventory system where $n = 2$ are examined. Table 3.7 summarizes the full factorial experimental design of the problem instances. Similar to the $n = 1$ case, the part failure distribution is assumed exponential, with a rate parameter $\lambda = 0.01$, or a mean time between part failures $\mu = 100$ for the two customers in the supply network.

Table 3.7. Ranges of the problem instance parameters for a two-echelon, single-supplier two-customer service parts inventory supply chain system.

	Supplier	Customer 1	Customer 2
Unit Ordering Cost, c_o	-	10	15
Unit Holding Cost, c_h	[300, 600]	[300, 600]	[300, 600]
Unit Replacement Cost, c_r	-	30	25
Unit Failure Cost, c_f	-	28	20
Unit Shortage Cost, c_s	-	[300, 600]	[300, 600]
Mean Time Between Failure, μ	-	100	100
Unit Emergency Shipping Cost, c_e	[300, 600]	-	-
Unit Production Cost, c_p	50	-	-
Unit Production Time, p	4	-	-
Order Delivery Lead Time, L		2	3

3.6.2.1. Decoupled Optimization Approach

In this section, we generate the joint inventory-maintenance policy under decoupled optimization. We use the parameter values and ranges listed in Table 3.7. Table 3.8 shows the specific parameter values used in this study. “S” in the table heading indicates supplier-related parameters and results and “C1” and “C2” indicate Customer 1-related parameters and results and Customer 2-related parameters and results, respectively. Table 3.9 summarizes the part replacement and part ordering policies for the customers at the different parameter levels, and the optimal part production schedule at the supplier for each customer.

Table 3.8. Specific problem instance cost parameters for a two-echelon, one supplier, two-customer service parts inventory supply chain system.

Prob Inst	c_p S	c_h^s S	c_e S	c_r C1	c_r C2	c_f C1	c_f C2	c_o C1	c_o C2	c_h^c C1	c_h^c C2	c_s C1	c_s C2
1	50	300	300	30	25	28	20	10	15	300	300	300	300
2	50	600	300	30	25	28	20	10	15	300	300	300	300
3	50	300	300	30	25	28	20	10	15	600	300	300	300
4	50	600	300	30	25	28	20	10	15	600	300	300	300
5	50	300	300	30	25	28	20	10	15	300	600	300	300
6	50	600	300	30	25	28	20	10	15	300	600	300	300
7	50	300	300	30	25	28	20	10	15	600	600	300	300
8	50	600	300	30	25	28	20	10	15	600	600	300	300
9	50	300	300	30	25	28	20	10	15	300	300	600	300
10	50	600	300	30	25	28	20	10	15	300	300	600	300
11	50	300	300	30	25	28	20	10	15	600	300	600	300
12	50	600	300	30	25	28	20	10	15	600	300	600	300
13	50	300	300	30	25	28	20	10	15	300	600	600	300
14	50	600	300	30	25	28	20	10	15	300	600	600	300
15	50	300	300	30	25	28	20	10	15	600	600	600	300
16	50	600	300	30	25	28	20	10	15	600	600	600	300
17	50	300	300	30	25	28	20	10	15	300	300	300	600
18	50	600	300	30	25	28	20	10	15	300	300	300	600
19	50	300	300	30	25	28	20	10	15	600	300	300	600
20	50	600	300	30	25	28	20	10	15	600	300	300	600
21	50	300	300	30	25	28	20	10	15	300	600	300	600
22	50	600	300	30	25	28	20	10	15	300	600	300	600
23	50	300	300	30	25	28	20	10	15	600	600	300	600
24	50	600	300	30	25	28	20	10	15	600	600	300	600
25	50	300	300	30	25	28	20	10	15	300	300	600	600
26	50	600	300	30	25	28	20	10	15	300	300	600	600
27	50	300	300	30	25	28	20	10	15	600	300	600	600
28	50	600	300	30	25	28	20	10	15	600	300	600	600
29	50	300	300	30	25	28	20	10	15	300	600	600	600
30	50	600	300	30	25	28	20	10	15	300	600	600	600
31	50	300	300	30	25	28	20	10	15	600	600	600	600
32	50	600	300	30	25	28	20	10	15	600	600	600	600
33	50	300	600	30	25	28	20	10	15	300	300	300	300
34	50	600	600	30	25	28	20	10	15	300	300	300	300
35	50	300	600	30	25	28	20	10	15	600	300	300	300
36	50	600	600	30	25	28	20	10	15	600	300	300	300
37	50	300	600	30	25	28	20	10	15	300	600	300	300
38	50	600	600	30	25	28	20	10	15	300	600	300	300
39	50	300	600	30	25	28	20	10	15	600	600	300	300
40	50	600	600	30	25	28	20	10	15	600	600	300	300
41	50	300	600	30	25	28	20	10	15	300	300	600	300
42	50	600	600	30	25	28	20	10	15	300	300	600	300

Table 3.8. (cont'd) Specific problem instance cost parameters for a two-echelon, one supplier, two-customer service parts inventory supply chain system.

Prob Inst	c_p S	c_h^s S	c_e S	c_r C1	c_r C2	c_f C1	c_f C2	c_o C1	c_o C2	c_h^c C1	c_h^c C2	c_s C1	c_s C2
43	50	300	600	30	25	28	20	10	15	600	300	600	300
44	50	600	600	30	25	28	20	10	15	600	300	600	300
45	50	300	600	30	25	28	20	10	15	300	600	600	300
46	50	600	600	30	25	28	20	10	15	300	600	600	300
47	50	300	600	30	25	28	20	10	15	600	600	600	300
48	50	600	600	30	25	28	20	10	15	600	600	600	300
49	50	300	600	30	25	28	20	10	15	300	300	300	600
50	50	600	600	30	25	28	20	10	15	300	300	300	600
51	50	300	600	30	25	28	20	10	15	600	300	300	600
52	50	600	600	30	25	28	20	10	15	600	300	300	600
53	50	300	600	30	25	28	20	10	15	300	600	300	600
54	50	600	600	30	25	28	20	10	15	300	600	300	600
55	50	300	600	30	25	28	20	10	15	600	600	300	600
56	50	600	600	30	25	28	20	10	15	600	600	300	600
57	50	300	600	30	25	28	20	10	15	300	300	600	600
58	50	600	600	30	25	28	20	10	15	300	300	600	600
59	50	300	600	30	25	28	20	10	15	600	300	600	600
60	50	600	600	30	25	28	20	10	15	600	300	600	600
61	50	300	600	30	25	28	20	10	15	300	600	600	600
62	50	600	600	30	25	28	20	10	15	300	600	600	600
63	50	300	600	30	25	28	20	10	15	600	600	600	600
64	50	600	600	30	25	28	20	10	15	600	600	600	600

Table 3.9. Service part ordering, part replacement, and part production times when each supply chain member's cost objective is optimized separately.

Prob Inst	t_o C1	t_r C1	t_p C1	t_o C2	t_r C2	t_p C2	Cost/ Time S	Cost/ Time C1	Cost/ Time C2	Long-Run Total Cost/ Time
1	4240.17	4530.69	1548.8	4240.17	4530.69	1527.2	25.29	7.14	9.71	42.14
2	4240.17	4530.69	1548.8	4240.17	4530.69	1527.2	25.29	7.14	9.71	42.14
3	4240.17	4530.69	264.8	4240.17	4530.69	1527.2	27.30	13.65	18.88	59.83
4	4240.17	4530.69	264.8	4240.17	4530.69	1527.2	27.30	13.65	18.88	59.83
5	4240.17	4530.69	1548.8	4240.17	4530.69	254.4	25.29	7.14	9.71	42.14
6	4240.17	4530.69	1548.8	4240.17	4530.69	254.4	25.29	7.14	9.71	42.14
7	4240.17	4530.69	264.8	4240.17	4530.69	254.4	25.29	7.14	9.71	42.14
8	4240.17	4530.69	264.8	4240.17	4530.69	254.4	25.29	7.14	9.71	42.14
9	4240.17	4530.69	4102.4	4240.17	4530.69	1527.2	25.29	13.65	9.71	48.65
10	4240.17	4530.69	4102.4	4240.17	4530.69	1527.2	25.29	13.65	9.71	48.65
11	4240.17	4530.69	1511.2	4240.17	4530.69	1527.2	25.29	13.65	9.71	48.65
12	4240.17	4530.69	1511.2	4240.17	4530.69	1527.2	25.29	13.65	9.71	48.65
13	4240.17	4530.69	4102.4	4240.17	4530.69	254.4	25.29	13.65	9.71	48.65
14	4240.17	4530.69	4102.4	4240.17	4530.69	254.4	25.29	13.65	9.71	48.65
15	4240.17	4530.69	1511.2	4240.17	4530.69	254.4	25.29	13.65	9.71	48.65
16	4240.17	4530.69	1511.2	4240.17	4530.69	254.4	25.29	13.65	9.71	48.65
17	4240.17	4530.69	1548.8	4240.17	4530.69	4080.8	25.29	7.14	18.88	51.31
18	4240.17	4530.69	1548.8	4240.17	4530.69	4080.8	25.29	7.14	18.88	51.31
19	4240.17	4530.69	264.8	4240.17	4530.69	4080.8	25.29	7.14	18.88	51.31
20	4240.17	4530.69	264.8	4240.17	4530.69	4080.8	25.29	7.14	18.88	51.31
21	4240.17	4530.69	1548.8	4240.17	4530.69	1500.8	25.29	7.14	18.88	51.31
22	4240.17	4530.69	1548.8	4240.17	4530.69	1500.8	25.29	7.14	18.88	51.31
23	4240.17	4530.69	264.8	4240.17	4530.69	1500.8	25.29	7.14	18.88	51.31
24	4240.17	4530.69	264.8	4240.17	4530.69	1500.8	25.29	7.14	18.88	51.31
25	4240.17	4530.69	4102.4	4240.17	4530.69	4080.8	25.29	13.65	18.88	57.82
26	4240.17	4530.69	4102.4	4240.17	4530.69	4080.8	25.29	13.65	18.88	57.82
27	4240.17	4530.69	1511.2	4240.17	4530.69	4080.8	25.29	13.65	18.88	57.82
28	4240.17	4530.69	1511.2	4240.17	4530.69	4080.8	25.29	13.65	18.88	57.82
29	4240.17	4530.69	4102.4	4240.17	4530.69	1500.8	25.29	13.65	18.88	57.82
30	4240.17	4530.69	4102.4	4240.17	4530.69	1500.8	25.29	13.65	18.88	57.82
31	4240.17	4530.69	1511.2	4240.17	4530.69	1500.8	25.29	13.65	18.88	57.82
32	4240.17	4530.69	1511.2	4240.17	4530.69	1500.8	25.29	13.65	18.88	57.82
33	4240.17	4530.69	1548.8	4240.17	4530.69	1527.2	27.30	7.14	9.71	44.14
34	4240.17	4530.69	1548.8	4240.17	4530.69	1527.2	27.29	7.14	9.71	44.14
35	4240.17	4530.69	264.8	4240.17	4530.69	1527.2	27.30	7.14	9.71	44.14
36	4240.17	4530.69	264.8	4240.17	4530.69	1527.2	27.29	7.14	9.71	44.14
37	4240.17	4530.69	1548.8	4240.17	4530.69	254.4	27.30	7.14	9.71	44.14
38	4240.17	4530.69	1548.8	4240.17	4530.69	254.4	27.29	7.14	9.71	44.14
39	4240.17	4530.69	264.8	4240.17	4530.69	254.4	27.30	7.14	9.71	44.14
40	4240.17	4530.69	264.8	4240.17	4530.69	254.4	27.29	7.14	9.71	44.14
41	4240.17	4530.69	4102.4	4240.17	4530.69	1527.2	27.30	13.65	9.71	50.66
42	4240.17	4530.69	4102.4	4240.17	4530.69	1527.2	27.29	13.65	9.71	50.66

Table 3.9 (cont'd) Service part ordering, part replacement, and part production times when each supply chain member's cost objective is optimized separately under the decoupled approach.

Prob Inst	t_o C1	t_r C1	t_p C1	t_o C2	t_r C2	t_p C2	Cost/ Time S	Cost/ Time C1	Cost/ Time C2	Long-Run Total Cost/ Time
43	4240.17	4530.69	1511.2	4240.17	4530.69	1527.2	27.30	13.65	9.71	50.66
44	4240.17	4530.69	1511.2	4240.17	4530.69	1527.2	27.29	13.65	9.71	50.66
45	4240.17	4530.69	4102.4	4240.17	4530.69	254.4	27.30	13.65	9.71	50.66
46	4240.17	4530.69	4102.4	4240.17	4530.69	254.4	27.29	13.65	9.71	50.66
47	4240.17	4530.69	1511.2	4240.17	4530.69	254.4	27.30	13.65	9.71	50.66
48	4240.17	4530.69	1511.2	4240.17	4530.69	254.4	27.29	13.65	9.71	50.66
49	4240.17	4530.69	1548.8	4240.17	4530.69	4080.8	27.30	7.14	18.88	53.31
50	4240.17	4530.69	1548.8	4240.17	4530.69	4080.8	27.29	7.14	18.88	53.31
51	4240.17	4530.69	264.8	4240.17	4530.69	4080.8	27.30	7.14	18.88	53.31
52	4240.17	4530.69	264.8	4240.17	4530.69	4080.8	27.29	7.14	18.88	53.31
53	4240.17	4530.69	1548.8	4240.17	4530.69	1500.8	27.30	7.14	18.88	53.31
54	4240.17	4530.69	1548.8	4240.17	4530.69	1500.8	27.29	7.14	18.88	53.31
55	4240.17	4530.69	264.8	4240.17	4530.69	1500.8	27.30	7.14	18.88	53.31
56	4240.17	4530.69	264.8	4240.17	4530.69	1500.8	27.29	7.14	18.88	53.31
57	4240.17	4530.69	4102.4	4240.17	4530.69	4080.8	27.30	13.65	18.88	59.83
58	4240.17	4530.69	4102.4	4240.17	4530.69	4080.8	27.29	13.65	18.88	59.82
59	4240.17	4530.69	1511.2	4240.17	4530.69	4080.8	27.30	13.65	18.88	59.83
60	4240.17	4530.69	1511.2	4240.17	4530.69	4080.8	27.29	13.65	18.88	59.82
61	4240.17	4530.69	4102.4	4240.17	4530.69	1500.8	27.30	13.65	18.88	59.83
62	4240.17	4530.69	4102.4	4240.17	4530.69	1500.8	27.29	13.65	18.88	59.82
63	4240.17	4530.69	1511.2	4240.17	4530.69	1500.8	27.30	13.65	18.88	59.83
64	4240.17	4530.69	1511.2	4240.17	4530.69	1500.8	27.29	13.65	18.88	59.82

3.6.2.2. Joint Optimization Approach

The search control parameters for the real-coded NSGA-II used in the single customer configuration are used in this section. These parameters are summarized in Table 3.5. Two representative problem instances are examined for convergence behavior. Figure 3.5 and Figure 3.6 show the Pareto efficient frontier for Customer 1 vs. Customer 2 for Prob Inst 30, and Customer 1 vs. Customer 2 for Prob Inst 38, respectively. Figure 3.7 shows the Customer 1 vs. Customer 2 vs. the Supplier for Prob Insts 30 and 38. We see that the Pareto front in the tri-objective space along the curve is convex. In addition, there is also a reasonable level of diversity among the solutions along the Pareto front.

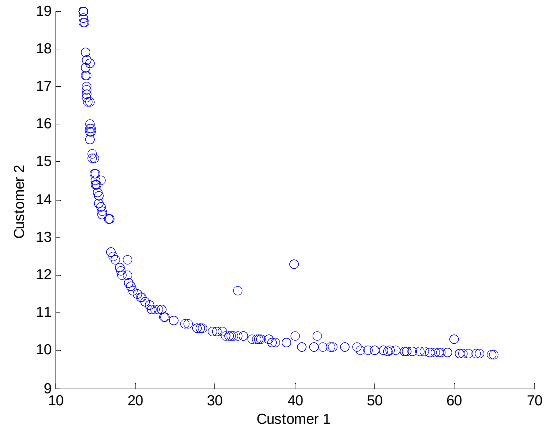

Figure 3.4. Customer 1 long-run expected total cost per unit time vs. Customer 2 long-run expected total cost per unit time for Prob Inst 30.

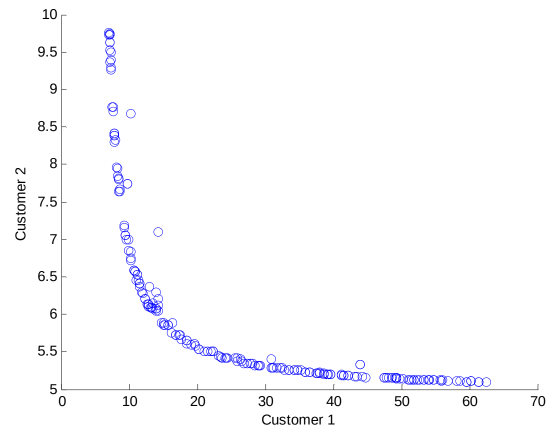

Figure 3.5. Customer 1 long-run expected total cost per unit time vs. Customer 2 long-run expected total cost per unit time for Prob Inst 38.

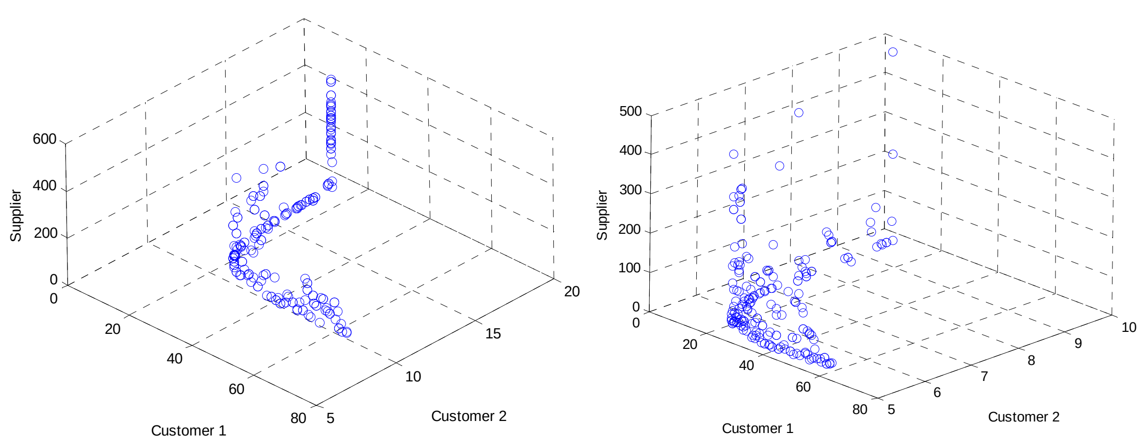

Figure 3.6. Customer 1 long-run expected total cost per unit time vs. Customer 2 long-run expected total cost per unit time vs. Supplier long-run expected total cost per unit time for Prob Insts 30 and 38.

NSGA-II generates a set of nondominated solutions for each of the 64 problem instances. Of the set of Pareto optimal solutions for a problem instance, at least one solution has the lowest long-run expected total cost per unit time. The t_o, t_r and t_p values that generated the lowest cost rate for each problem instance are also reported.

In order to make a reasonable comparison of these results, the same set of problem parameter values are used for the decoupled optimization, as shown in Table 3.8. The results under joint optimization are summarized in Table 3.10. The average improvement over decoupled optimization over the 64 problem instances in 48.98% with a standard deviation of 7.68% ranging from 18.85% (Prob Insts 3 and 4) to 59.63% (Prob Insts 33 to 40). These results show that there is economic benefit to simultaneously optimizing objectives of the individual members of the supply chain network. Furthermore, it is important to note that the strength of the multiobjective optimization approach is that it can handle creating a set of alternatives in (n+1) dimensional decision space.

Table 3.10. Service part ordering, part replacement and part production times when each supply chain member's cost objective is optimized simultaneously under the joint approach.

Prob Inst	t_o C1	t_r C1	t_p C1	t_o C2	t_r C2	t_p C2	Cost/ Time S	Cost/ Time C1	Cost/ Time C2	Long-Run Total Cost/ Time	% Improve over Decoupled
1	3968.52	4496.05	3963.33	4567.19	4951.72	4384.41	0.99	8.10	10.68	19.77	53.08%
2	4091.84	4936.59	4083.99	4561.29	4977.24	4399.89	0.99	8.10	10.68	19.77	53.08%
3	4383.19	4657.34	4315.14	4561.46	4895.38	4380.19	0.99	23.04	24.46	48.49	18.95%
4	4208.27	4464.67	4195.30	4510.99	4581.55	4380.63	0.99	23.04	24.46	48.49	18.95%
5	4627.74	4909.43	4428.27	4443.59	4913.58	4310.66	1.00	7.06	9.76	17.82	57.71%
6	4314.16	4711.41	4293.77	4625.64	4834.97	4408.82	1.00	7.06	9.76	17.82	57.71%
7	4607.75	4989.49	4491.91	4439.44	4974.40	4260.05	1.00	7.06	9.76	17.82	57.71%
8	4315.12	4576.55	4281.44	3963.47	3981.58	3961.33	1.00	7.06	9.76	17.82	57.71%
9	3947.03	3975.98	3945.53	4344.89	4816.02	4295.83	1.00	13.46	9.76	24.23	50.20%
10	4334.11	4986.44	4281.43	4276.05	4988.62	4270.46	1.00	13.46	9.76	24.23	50.20%
11	4323.56	4449.79	4265.13	4243.88	4811.26	4210.73	1.00	13.46	9.76	24.23	50.20%
12	4453.22	4964.36	4394.69	4697.60	4844.12	4478.85	1.00	13.46	9.76	24.23	50.20%
13	4680.69	4984.36	4518.91	4567.42	4684.61	4346.02	1.00	13.46	9.76	24.23	50.20%
14	4098.65	4756.71	4092.77	4624.95	4961.47	4492.30	1.00	13.46	9.76	24.23	50.20%
15	4539.19	4836.51	4522.65	4368.08	4984.67	4289.30	1.00	13.46	9.76	24.23	50.20%
16	4281.71	4946.53	4233.30	4666.00	4944.44	4407.91	1.00	13.46	9.76	24.23	50.20%
17	4455.90	4671.18	4409.73	4445.30	4936.92	4251.60	1.00	7.06	18.96	27.02	47.34%
18	4146.25	4483.86	4142.14	4573.71	4978.09	4564.81	1.00	7.06	18.96	27.02	47.34%
19	4404.52	4831.77	4328.28	4693.69	4758.35	4364.72	1.00	7.06	18.96	27.02	47.34%
20	4312.26	4851.13	4249.34	4539.64	4688.57	4437.02	1.00	7.06	18.96	27.02	47.34%
21	4621.97	4961.65	4612.54	4150.94	4430.03	4135.39	1.00	7.06	18.96	27.02	47.34%
22	4205.63	4590.33	4187.98	4747.08	4999.49	4447.83	1.00	7.06	18.96	27.02	47.34%
23	4695.76	4964.83	4368.22	4660.77	4840.08	4635.66	1.00	7.06	18.96	27.02	47.34%
24	4354.23	4861.81	4258.30	4706.83	4981.78	4538.96	1.00	7.06	18.96	27.02	47.34%
25	4634.28	4931.65	4546.93	3766.13	3785.84	3764.75	1.00	13.46	18.96	33.42	42.20%
26	4354.93	4969.11	4322.63	4619.75	4920.80	4481.41	1.00	13.46	18.96	33.42	42.20%
27	4669.36	4951.02	4408.01	4714.34	4956.10	4692.43	1.00	13.46	18.96	33.42	42.20%
28	4642.56	4986.07	4412.92	4108.80	4806.33	4103.20	1.00	13.46	18.96	33.42	42.20%
29	4011.39	4967.85	4002.50	4705.41	4878.99	4457.76	1.00	13.46	18.96	33.42	42.20%
30	4414.95	4986.96	4383.56	4536.36	4977.07	4313.97	1.00	13.46	18.96	33.42	42.20%
31	4687.57	4753.39	4564.76	4361.77	4952.16	4301.45	1.00	13.46	18.96	33.42	42.20%
32	4720.60	4983.08	4644.43	3972.64	4965.16	3968.68	1.00	13.46	18.96	33.42	42.20%
33	4591.34	4876.98	4497.14	4450.29	4871.80	4306.07	1.00	7.06	9.76	17.82	59.63%
34	4475.75	4860.27	4370.83	4497.12	4933.08	4389.96	1.00	7.06	9.76	17.82	59.63%
35	4531.64	4883.39	4340.32	4086.92	4926.31	4078.57	1.00	7.06	9.76	17.82	59.63%
36	4631.63	4987.69	4442.76	4092.09	4795.26	4089.97	1.00	7.06	9.76	17.82	59.63%
37	4660.13	4979.17	4560.09	4391.70	4906.07	4321.37	1.00	7.06	9.76	17.82	59.63%
38	4459.03	4856.67	4411.31	4546.02	4795.26	4469.72	1.00	7.06	9.76	17.82	59.63%
39	4635.80	4833.03	4491.39	4355.68	4514.62	4300.46	1.00	7.06	9.76	17.82	59.63%
40	4413.77	4483.39	4281.29	4589.96	4862.22	4437.41	1.00	7.06	9.76	17.82	59.63%
41	3789.02	3796.95	3787.03	4367.40	4991.20	4307.18	1.00	13.46	9.76	24.23	52.17%
42	4131.51	4731.49	4129.74	4432.44	4657.06	4426.25	1.00	13.46	9.76	24.23	52.17%

Table 3.10 (cont'd) Service part ordering, part replacement and part production times when each supply chain member's cost objective is optimized simultaneously.

Prob Inst	t_o C1	t_r C1	t_p C1	t_o C2	t_r C2	t_p C2	Cost/ Time S	Cost/ Time C1	Cost/ Time C2	Long-Run Total Cost/ Time	% Improve over Decoupled
43	4246.25	4395.27	4236.57	4302.89	4687.18	4274.25	1.00	13.46	9.76	24.23	52.17%
44	4166.03	4292.23	4157.74	4661.85	4897.28	4427.99	1.00	13.46	9.76	24.23	52.17%
45	4192.65	4902.69	4138.41	4179.04	4923.47	4165.41	1.00	13.46	9.76	24.23	52.17%
46	4018.03	4584.17	4014.99	4149.42	4500.49	4137.39	1.00	13.46	9.76	24.23	52.17%
47	4854.19	4986.25	4616.24	4285.68	4973.48	4265.96	1.00	13.46	9.76	24.23	52.17%
48	4093.99	4931.86	4089.55	4262.78	4490.70	4228.30	1.00	13.46	9.76	24.23	52.17%
49	3922.98	3947.00	3920.50	4460.95	4955.14	4250.26	1.00	7.06	18.96	27.02	49.32%
50	4095.07	4638.69	4092.87	3910.74	4985.03	3909.13	1.00	7.06	18.96	27.02	49.31%
51	4391.05	4875.53	4244.91	4555.21	4941.00	4335.50	1.00	7.06	18.96	27.02	49.32%
52	4282.07	4880.14	4221.59	4524.00	4564.43	4401.11	1.00	7.06	18.96	27.02	49.31%
53	4864.71	4975.87	4367.20	4423.23	4894.25	4219.37	1.00	7.06	18.96	27.02	49.32%
54	4466.71	4747.99	4296.54	4472.85	4974.35	4440.75	1.00	7.06	18.96	27.02	49.31%
55	4198.03	4309.78	4190.31	4485.54	4937.59	4372.17	1.00	7.06	18.96	27.02	49.32%
56	4723.01	4978.41	4702.08	4226.60	4755.43	4220.89	1.00	7.06	18.96	27.02	49.31%
57	3921.68	3965.58	3919.64	4007.07	4564.65	4004.60	1.00	13.46	18.96	33.42	44.14%
58	4346.52	4857.56	4315.37	3903.72	4762.75	3902.05	1.00	13.46	18.96	33.42	44.14%
59	4104.26	4773.09	4095.41	4005.19	4639.01	4001.31	1.00	13.46	18.96	33.42	44.14%
60	4736.29	4951.11	4569.47	4058.80	4907.77	4056.28	1.00	13.46	18.96	33.42	44.14%
61	4011.81	4287.94	4010.38	4430.69	4961.70	4322.83	1.00	13.46	18.96	33.42	44.14%
62	4220.92	4342.94	4188.48	4116.17	4444.22	4111.51	1.00	13.46	18.96	33.42	44.14%
63	4692.94	4906.78	4369.60	4491.29	4952.34	4421.27	1.00	13.46	18.96	33.42	44.14%
64	4505.75	4966.29	4284.66	3999.20	4772.80	3996.32	1.00	13.46	18.96	33.42	44.14%

3.6.3. Equitable Apportionment of the Economic Benefit of Simultaneous Optimization

The ultimate goal of this multiobjective modeling approach for service parts inventory and maintenance is to benefit all members involved in the supply chain system (*i.e.*, the parts supplier and the *n* customers). For that reason, it is important to determine the best way in which all the members are compensated appropriately for what they are gaining (or sacrificing) to generate the overall system-wide cost savings. Usually, this is determined by a central decision-maker who attempts to find a balance so that all members of the supply chain are fairly treated.

As previously discussed, LTSAs are formal vehicles offered by OEMs to help their customers (equipment owners and/or operators) maximize the availability of their equipment. These agreements place the responsibility of planned maintenance scheduling and service parts availability supporting a set (or pool) of equipment on the OEM rather than the equipment owner. A major challenge with LTSAs is negotiating the terms under which the customer and the part supplier will partner. These negotiations should include the equitable allocation of all cost savings achieved under the agreement. Table 3.11 shows an example allocation policy. Using Prob Inst 34 that yields the greatest percent improvement in system-wide costs, the computations are summarized in the table. Notice that Customer 2 must sacrifice in terms of cost savings but the Supplier and Customer 1 experience cost savings – 93.89% and 8.16%, respectively. Overall, the system-wide cost savings is 43.74%.

Table 3.11. Comparison of costs per unit time for the supplier and two customers (using Prob Inst 34).

Optimization Approach	t_o C1	t_r C1	t_p C1	t_o C2	t_r C2	t_p C2	Cost/ Time S	Cost/ Time C1	Cost/ Time C2	Long-Run Total Cost/ Time
Decoupled	4240.17	4530.69	1548.8	4240.17	4530.69	1527.2	27.29	7.14	9.71	44.14
Joint	4475.75	4860.27	4370.83	4497.12	4933.08	4389.96	1.00	7.06	9.76	17.82
Savings							26.29	0.08	-0.05	26.32
% Improve							96.34%	1.12%	-0.51%	59.63%

In the case of multi-echelon service parts supply chain systems, customers often enter into service parts inventory pools and pay a fee that warranties their supplies. For instance, the decision-maker can use the potential savings (or sacrifice) and the proportion of savings (or sacrifice) that a customer experiences in the partnership to determine the membership fee that should be charged for that customer or possible discounts in any of the services or the part itself. The main idea is that everybody in the system should benefit from this agreement. We illustrate two possible alternatives that can be used for this purpose.

A first, and perhaps somewhat simple, alternative considers only the overall expected cost savings of the system. In this case, the individual savings are not used for the calculations.

All participants, independently of the amount of savings, will have the same saving percentage. Using Prob Inst 34, each entity in the system will have potential savings of 59.63% as shown in Table 3.11. Based on these long-run expected cost rates, the decision-maker can calculate a base membership fee for each customer or a discount as you can see in Table 3.12. Although it has been demonstrated that multiobjective optimization works better than individual optimization, it might be the case in which a customer does not benefit from the decoupled optimization. In that situation, it makes little economic sense for anyone to have that customer in the system, since the overall costs are going to be affected and the supplier would have to sacrifice or incur additional costs (*e.g.*, decreasing the price of the service part) so that customer can enter the agreement.

Table 3.12 Summary of apportioning alternatives.

Optimization Approach	Cost/Time S	Cost/Time C1	Cost/Time C2	Long-Run Total Cost/ Time
Individual	27.29	7.14	9.71	44.14
Joint	1.00	7.06	9.76	17.82
			Overall Cost Savings	26.32
			Overall % Improve Savings	59.63%
Individual Costs	16.27	4.26	5.79	
Base Membership Fee/Discount		2.88	3.92	

A second alternative could be based on the opportunity cost (OC) of the supplier, in that, it is based on the profit the supplier expects to realize given the potential savings using joint optimization. To illustrate using the information in Table 3.11, we assume that the OC of the parts supplier is 35%. The supplier expected savings are \$27.29 – (0.35)(\$27.29) = \$9.55. In order to realize these savings, the supplier should give up \$26.29 – \$9.55 = \$16.74 that should be proportionally distributed to the customers (n = 2, in this example). Although these examples are somewhat simple, we note here that it is, in fact, the onus of the negotiators to decide on how the potential savings should be distributed and how the fees are determined.

3.7. Summary and Conclusions

In this chapter, we develop the cost models for a joint inventory and maintenance service parts inventory system for a two-echelon, one supplier, *n*-customer configuration. Next, we use individual and simultaneous optimization algorithms (RCGA and NSGA-II, respectively) to determine the best ordering and replacement policies that minimize the system-wide combined long-run expected total cost per unit time. We perform experiments varying the number of customers served by the supplier to show the difference between the two approaches.

According to the results, we conclude that the simultaneous optimization approach outperforms the individual optimization approach. In the experiment with only one customer in the system, the system-wide cost per unit time is improves in most of the problem instances, and the average improvement is 43.52%. Similarly, in the experiment with two customers in the system, the benefits of using a multiobjective approach are also shown, where the average improvement is shown to be 59.63%.

In the next chapter, the assumption of no lateral transshipments is relaxed. In other words, we extend the analysis to a model that allows lateral transshipments between the customers. Later, we compare the optimization of the both supply chain configurations, with and without lateral transshipments.

CHAPTER 4:
JOINT MAINTENANCE AND SERVICE PARTS INVENTORY MULTIOBJECTIVE OPTIMIZATION FOR A TWO-ECHELON SINGLE SUPPLIER AND *n*-CUSTOMER SUPPLY CHAIN SYSTEM WITH LATERAL TRANSSHIPMENTS

4.1. Introduction

In this chapter, an extension of the joint maintenance and service parts inventory model for a two-echelon, single supplier, n-customer supply chain system is explored. The assumption of no lateral transshipments is relaxed as depicted in Figure 4.1. First, we explain the details of the model, its parameters and the relevant decision variables. After this, long-run expected cost functions are each derived for the customers and the supplier. Finally, experiments are performed to compare the impact of using lateral transshipments versus no lateral transshipments.

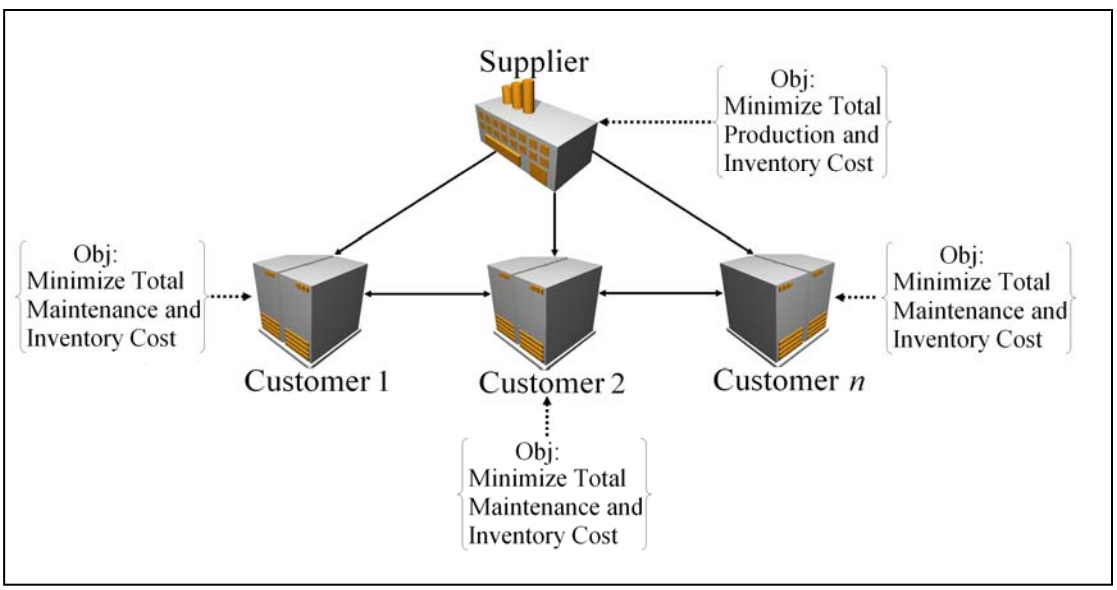

Figure 4.1. The two-echelon, single-supplier, n-customer inventory system configuration with lateral transshipments.

4.2. Description of the Generalized Multi-Echelon Service Parts Supply Chain System with Lateral Transshipments

Lateral transshipments (LTs) are used with the purpose of covering more expansive areas geographically compared to the area that a single supplier would be able to cover. Previous research has shown that LTs help to reduce response times and that they are economically beneficial, not only in terms of reduced downtimes, but due to lower shipping costs and higher availability of machines (Lee, 1987; Axsäter, 1990b; Archibald *et al.*, 1997; Alfredsson and Verrijdt, 1999; Grahovac and Chakravarty, 2001; Wong *et al.*, 2006).

In a service parts pooling system that uses lateral transshipments, the primary goal is to reduce and balance system-wide costs. For that reason, there are tradeoffs that partnering customers must consider in order to benefit the entire supply system. For instance, it might be the case that, for a specific customer, utilizing more LTs is more economically beneficial than using regular shipments from the supplier due to high holding costs or penalties. This could negatively impact the overall savings of the system. However, when the decisions are made for the benefit of all members of the supply system, the total savings can outweigh the individual member benefits.

Here, we consider a similar system as the one described in the previous chapter. This system consists of a single service part, one supplier, and n customers (see Figure 4.1). Like the previous configuration, the customers are individual systems with their own objectives to satisfy. Customers want to minimize their long-run expected total maintenance and service parts inventory costs, while the supplier wants to minimize its long-run expected total production and inventory costs. Customers must determine the best source for inventory replenishment (*i.e.*, the supplier or another customer within the supply system). Furthermore, customers must decide the best time to order a part and the best time to replace a part in order to minimize their expected total maintenance and service parts inventory costs. Alternatively, the supplier must decide the best time to produce service parts in order to satisfy the customers' demand for them and minimize its production and inventory costs. Before presenting the formulations, we give the relevant notation, parameters, variables and modeling assumptions.

Relevant Notation, Parameters and Variables:

$f(t)$: Probability density function of the service part failure distribution as a function of t.

Customer-Specific Parameters:

L : Lead time for a customer to receive an ordered service part from the supplier; (unit of time, *e.g.*, minutes, hours, days, *etc.*)

L_i : Lead time for a customer to receive an ordered service part from customer i; (unit of time, *e.g.*, minutes, hours, days, *etc.*)

c_r : Part replacement cost at a customer; represents all costs incurred when a part is replaced including the cost of the part and the labor cost; (cost per replacement)

c_f : Part failure cost at a customer; represents the costs associated with a failure including repairing any damage that may have occurred at the time of failure and any subsequent damage to the system until the failed part is replaced; (cost per failure)

c_s : Part shortage cost at a customer; (cost per unit time)

c_h^c : Part holding cost at a customer; (cost per unit time)

c_o : Part ordering cost from the supplier; represents all costs associated with placing an order and when that order is received, including any clerical/labor costs of processing orders, inspection and return of poor quality units and material handling costs; (cost per order)

c_o^i : Part ordering cost from customer i; it includes the same aspects considered for c_o.

Supplier-Specific Parameters:

$g(n)$: Probability mass function of the supplier having a service part in stock as a function of n, the number of customers serviced by the supplier

c_h^s : Part holding cost at the supplier; (cost per unit time)

c_p : Part production cost at the supplier; represents all costs associated with producing a part at the supplier including any production setup cost, or product; (cost per part)

p : Unit production time at the supplier; represents the production lead time if the supplier produces the part in-house or the delivery lead time of the part to the service part supplier from the supplier's supplier if production of the part is outsourced; (unit of time, *e.g.*, minutes, hours, days, *etc.*)

c_e : Emergency shipping cost incurred by the supplier; (cost per unit time)

Decision Variables:

t_o : Scheduled time to order a replacement service part from the supplier (at the customer)

t_o^i : Scheduled time to order a service part from partner customer i (at the customer)

t_r : Scheduled time to replace a service part (at the customer)

t_p : Scheduled time to produce a service part (at the supplier)

Modeling and Analysis Assumptions:

- Lateral transshipments of service parts between customers are allowed;
- Only one service part type is considered within the inventory supply system;
- The probability the supplier will have a part available when requested is a monotonically decreasing function of the number of customers;
- The parts are non-repairable (*i.e.*, replacement);
- Only one part is ordered at a time and a maximum of one service part is held in inventory, *i.e.*, base stock model;
- If the customer has service part inventory on-hand, then a new order cannot be placed by the customer;
- The time between successive part replacements is considered a cycle for a customer;
- The time between successive part production runs is considered a cycle for the supplier;
- Lead time L to receive a replacement part from the supplier after an order is placed is positive and fixed;
- Lead time L_i to receive a replacement part from another customer after an order is placed is positive and fixed;
- Actual part replacement once the part is received after L or L_i is assumed to be instantaneous and perfect;
- If a part is ordered from the supplier, the time at which a part is replaced cannot occur before the part arrives from the supplier, *i.e.*, $t_o + L + p \leq t_r$;
- If a part is ordered from another customer, the time at which a part is replaced cannot occur before the part arrives, *i.e.*, $t_o^i + L_i \leq t_r$;

- If a customer does not have a part in stock, orders from other customers cannot be received.
- No failure occurs during LTs since the lead time is considerably smaller than the mean time between failures;
- The different options that each customer has to replenish inventories are mutually exclusive; therefore, whatever source is best for the system will cancel out the variables associated with the other available sources;
- All activities required to expedite a customer order are assumed by the supplier; and
- All the costs are positive.

4.3. Modeling and Solving the Customer's Problem

The first portion of the model corresponds to the long-run expected cost function for a customer in the service part supply network. The long-run expected total cost per unit time for each customer includes the expected costs due to maintenance, the ordering of a service part, the inventory holding, and the inventory shortage, along with the expected cycle length. The most important addition to this model is the inclusion of lateral transshipments. This means that replenishments can be obtained from different sources (supplier or other customers), but only one source can be chosen. Since they are mutually exclusive, each option is represented as a single objective. For the optimization, the option that realizes the lower expected total cost per unit time is selected and the other options are discarded. Consequently, the customer can select between the options. The first option is requesting the part from the supplier. In this case, $E[TC_s]$ represents the expected total cost. The second option is requesting the part for a partner customer, and $E[TC_{ci}]$ represents the expected total cost. At this point, it is important to mention that an order cannot be placed to a partner customer unless a part in available. Thus, the long-run expected total cost is multiplied by the probability of the other customer having that part in stock. This is a function of that customer's time to replace t_r and its part failure probability distribution, which is described by $(1 - f(t_r))$. The objective functions are expressed as

$$E\left[TC_{cs}\right] = \frac{\begin{array}{l}\text{Ordering Cost} + E[\text{Maintenance Cost}] \ + \ E[\text{Inventory Holding Cost}] \\ + \ E[\text{Inventory Shortage Cost}]\end{array}}{E[\text{Cycle Length}]} \quad \text{and} \tag{4.1}$$

$$E\left[TC_{ci}\right] = \left(1 - f\left(t_r\right)\right)\left(\frac{\begin{array}{l}\text{Ordering Cost} + E[\text{Maintenance Cost}_i] \ + \ E[\text{Inventory Holding Cost}_i] \\ + \ E[\text{Inventory Shortage Cost}_i]\end{array}}{E[\text{Cycle Length}_i]}\right). \tag{4.2}$$

The expected maintenance cost is the sum of the unit replacement cost, c_r, and the expected cost due to a part failure before the part is replaced. Thus,

$$E\left[\text{Maintenance Cost}\right] = c_r \ + \ c_f \int_0^{t_r} f\left(t\right)\, dt \tag{4.3}$$

Holding cost is incurred if a replacement part is ordered and arrives before the existing part fails, from the time the replacement part arrives until it is used for preventive or corrective maintenance. As mentioned before, customers have two options to decide the source of replenishment. Consequently shortage costs functions for each case are developed. The resulting expected cost functions are

$$E\left[\text{Holding Cost}\right] = c_h^c \left[\begin{array}{l} g\left(n\right)\left(\displaystyle\int_{t_o+L}^{t_r} \left(t - t_o - L\right) f\left(t\right) dt + (t_r - t_o - L)\int_{t_r}^{\infty} f\left(t\right) dt\right) + \\ \left(1 - g\left(n\right)\right)\left(\displaystyle\int_{t_o+L+p}^{t_r} \left(t - t_o - L - p\right) f\left(t\right) dt + \ (t_r - t_o - L - p)\int_{t_r}^{\infty} f\left(t\right) dt\right)\end{array}\right] \quad \text{and} \tag{4.4}$$

$$E\left[\text{Holding Cost}_i\right] = c_h^c \left(\int_{t_{o_i}+L_i}^{t_r} \left(t - t_{o_i} - L_i\right) f\left(t\right) dt + (t_r - t_{o_i} - L_i)\int_{t_r}^{\infty} f\left(t\right) dt\right). \tag{4.5}$$

The shortage cost is incurred if a replacement part is ordered after an existing part fails until the part is received, either from the supplier or from another customer in the system. Shortage costs are also incurred if an existing part fails after the replacement part's order has been placed to the supplier or another customer, but before it arrives at the customer location. Finally, shortages costs can continue to accrue for the additional time associated with production (and/or procurement) if the supplier does not have a service part on hand to ship to the customer after the customer's order is received (E[Shortage Cost]) or different lead times if the order has been placed as a lateral transshipment (E[Shortage Cost$_i$]). Therefore, the customer's shortage cost can be expressed as

$$E\left[\text{Shortage Cost}\right] = c_s \left[\begin{array}{l} g(n)\left(\int_0^{t_o} Lf(t)\,dt + \int_{t_o}^{t_o+L} (t_o + L - t)\, f(t)\,dt \right) + \\ \left(1 - g(n)\right)\left(\int_0^{t_o} (L+p)\, f(t)\,dt + \int_{t_o}^{t_o+L+p} (t_o + L + p - t)\, f(t)\,dt \right) \end{array} \right] \text{ and} \tag{4.6}$$

$$E\left[\text{Shortage Cost}_i\right] = c_s \left(\int_0^{t_{o_i}} L_i\, f(t)\,dt + \int_{t_{o_i}}^{t_{o_i}+L_i} \left(t_{o_i} + L_i - t\right) f(t)\,dt \right). \tag{4.7}$$

As in the case where LTs are not allowed, the objective is to minimize the long-run expected total cost per unit time. The system regenerates every time a replacement takes place, thus the long-run cost rate is obtained by analyzing a cycle in the system. Then, the long-run expected cost rate is equal to the sum of the expected costs divided by the expected cycle length. Therefore, the joint expected cycle time for maintenance and inventories with lateral transshipments also has two parts that are incurred depending upon the best replenishment option as follows, Eqs. 4.8 and 4.9.

$$E[\text{Cycle Length}] = \begin{bmatrix} g(n)\left(\int_0^{t_o}(t+L)\,f(t)dt + \int_{t_o}^{t_o+L}(t_o+L)\,f(t)dt + \int_{t_o+L}^{t_r} tf(t)dt \; + \int_{t_r}^{\infty} t_r f(t)dt\right) \; + \\ (1-g(n))\left(\int_0^{t_o}(t+L+p)\,f(t)dt + \int_{t_o}^{t_o+L+p}(t_o+L+p)\,f(t)dt + \int_{t_o+L+p}^{t_r} tf(t)dt + \int_{t_r}^{\infty} t_r f(t)dt\right) \end{bmatrix} \tag{4.8}$$

and

$$E[\text{Cycle Length}_i] = \left(\int_0^{t_{o_i}}(t+L_i)\,f(t)dt + \int_{t_{o_i}}^{t_{o_i}+L_i}\left(t_{o_i}+L_i\right)f(t)dt + \int_{t_{o_i}+L_2}^{t_r} tf(t)dt \; + \int_{t_r}^{\infty} t_r f(t)\; dt\right). \tag{4.9}$$

Finally, we obtain the customer's long-run expected total cost per unit time by substituting Eqs. 4.3, 4.4, 4.6, and 4.8 into Eq. 4.1 and Eqs. 4.3, 4.5, 4.7, and 4.9 into Eq. 4.2.

$$E[TC_{cs}] = \frac{\begin{bmatrix} c_o + c_r + \; c_f \int_0^{t_r} f(t)dt + \\ c_h^c g(n)\left(\int_{t_o+L}^{t_r}(t-t_o-L)\,f(t)\;dt \; + \; (t_r-t_o-L)\int_{t_r}^{\infty} f(t)dt\right) + \\ c_h^c(1-g(n))\left(\int_{t_o+L+p}^{t_r}(t-t_o-L-p)\,f(t)\;dt + (t_r-t_o-L-p)\int_{t_r}^{\infty} f(t)dt\right) \\ c_s g(n)\left[\int_0^{t_o} Lf(t)\;dt + \int_{t_o}^{t_o+L}(t_o+L-t)\,f(t)dt\right] \; + \\ c_s(1-g(n))\left[\int_0^{t_o}(L+p)\,f(t)dt + \int_{t_o}^{t_o+L+p}(t_o+L+p-t)\,f(t)dt\right] \end{bmatrix}}{\begin{bmatrix} g(n)\left[\int_0^{t_o}(t+L)\,f(t)dt + \int_{t_o}^{t_o+L}(t_o+L)\,f(t)dt + \int_{t_o+L}^{t_r} tf(t)\;dt + \int_{t_r}^{\infty} t_r\; f(t)dt\right] \; + \\ (1-g(n))\left[\int_0^{t_o}(t+L+p)\,f(t)dt \; + \int_{t_o}^{t_o+L+p}(t_o+L+p)\,f(t)dt + \int_{t_o+L+p}^{t_r} tf(t)\;dt \; + \int_{t_r}^{\infty} t_r\; f(t)\;dt\right] \end{bmatrix}} \tag{4.10}$$

and

$$E\left[TC_{ci}\right]=\left(1-f\left(t_r\right)\right)\frac{\begin{array}{l} c_o^i+c_r+c_f\displaystyle\int_0^{t_r} f\left(t\right)dt+ \\ c_h^c\left(\displaystyle\int_{t_{o_i}+L_i}^{t_r}\left(t-t_{o_i}-L_i\right)f\left(t\right)dt+\left(t_r-t_{o_i}-L_i\right)\int_{t_r}^{\infty} f\left(t\right)dt\right)+ \\ c_s\left(\displaystyle\int_0^{t_{o_i}} L_i f\left(t\right)dt+\int_{t_{o_i}}^{t_{o_i}+L_i}\left(t_{o_i}+L_i-t\right)f\left(t\right)dt\right)\end{array}}{\left[\left(\displaystyle\int_0^{t_{o_i}}\left(t+L_i\right)f\left(t\right)dt+\int_{t_{o_i}}^{t_{o_i}+L_i}\left(t_{o_i}+L_i\right)f\left(t\right)dt+\int_{t_{o_i}+L_2}^{t_r} tf\left(t\right)dt+\int_{t_r}^{\infty} t_r f\left(t\right)dt\right)\right]}. \tag{4.11}$$

<u>Derivation of the Customer's Inventory and Maintenance Policies</u>

We assume that the random part failure follows a stationary exponential distribution. Given this, substituting the exponential probability density function into Eqs. 4.10 and 4.11, the long-run expected total cost per unit time is

$$E\left[TC_{cs}\right]=\frac{\begin{array}{l} c_r+c_f+c_f\,e^{-\lambda t_r}+c_o+ \\ c_h^c\left(\dfrac{e^{-\lambda\left(2L+p+2t_o+t_r\right)}\left(e^{\lambda\left(L+p+t_o+t_r\right)}g\left(n\right)-e^{\lambda\left(2L+p+2t_o\right)}-e^{\lambda\left(L+t_o+t_r\right)}\left(g\left(n\right)-1\right)\right)}{\lambda}\right)+ \\ c_s\left(L+p-g\left(n\right)p-\dfrac{e^{-\lambda\left(L+p+t_o\right)}\left(e^{\lambda\left(L+p\right)}+g\left(n\right)-e^{\lambda p}\left(g\left(n\right)-1\right)\right)}{\lambda}\right)\end{array}}{L+p-g\left(n\right)p-\dfrac{e^{-\lambda t_o}+e^{-\lambda t_r}+e^{-\lambda\left(L+p+t_o\right)}\left(g\left(n\right)-1\right)-e^{-\lambda\left(L+t_o\right)}\left(g\left(n\right)-1\right)}{\lambda}} \tag{4.12}$$

and

$$E\left[TC_{ci}\right]=\frac{\left(e^{-\lambda L_i}\left(e^{\lambda t_r}-\lambda\right)\left(-c_h^c-c_s+c_s e^{\lambda L_i}-c_s e^{-\lambda\left(L_i+t_{o_i}\right)}L_i\lambda+e^{\lambda\left(L_i-t_r+t_{o_i}\right)}\left(c_h^c+\left(c_f-e^{\lambda t_r}\left(c_{o_i}+c_r+c_f\right)\right)\lambda\right)\right)\right)}{\left(e^{\lambda t_{o_i}}+e^{\lambda t_r}-e^{-\lambda\left(L_i+t_r\right)}-e^{\lambda\left(t_r+t_{o_i}\right)}\left(1+L_i\lambda\right)\right)}. \tag{4.13}$$

There are two different scenarios that might occur with this configuration. The first one would be if both customers order the part from the supplier, then the derivations are exactly the same as in the model with no LTs. Then, t_{o_i} is not considered and t_o^* and t_r^* are equal to Eqs. 3.11

and 3.12. The other case takes place when a customer orders a part from another customer. In this scenario, the time to order for the customer is represented by t_{o_i}. To determine the optimal values of t_{o_i} and t_r, the partial derivatives $\partial E\left[TC_{ci}\right]/\partial t_{o_i}$ and $\partial E\left[TC_{ci}\right]/\partial t_r$ are set equal to zero and the resulting equations are solved simultaneously. It follows that the optimal values are

$$t_{o_i}^* = \frac{\ln\left(\dfrac{e^{-\lambda L_i}\left(c_h^c + c_s - c_s e^{\lambda L_i}\right)\lambda}{c_h^c + \lambda\left(c_f - \lambda\left(c_f + c_{o_i} + c_r - c_s L_i\right)\right)}\right)}{\lambda}, \text{ and} \tag{4.14}$$

$$t_r^* = \frac{\ln\left(\dfrac{c_h^c + c_f\lambda}{\lambda\left(c_f + c_{o_i} + c_r + c_s L_i\right)}\right)}{\lambda}. \tag{4.15}$$

It can be said that the total cost functions are pseudoconvex in t_r. Thus, for a given t_{o_i} the optimal t_r can be either ($t_{o_i} + L_i$), ∞ or t_r^*. In the case that t_r^* does not exist, the partial derivative $\partial E\left[TC_{ci}\right]/\partial t_r$ determines when to the part should be replaced. If it is negative, then the part should be replaced at failure. If it is positive, the part should be replaced just when the part is received from the other customer at ($t_{o_i} + L_i$).

Similar to the non-LTs model, it cannot be shown that the function is jointly convex on both variables using the determinant of the Hessian matrix. However, we can say that the total cost function is pseudoconvex in t_{o_i}. So, for a given t_r, the optimal t_{o_i} can either be 0, ($t_r - L_i$) or $t_{o_i}^*$. If $t_{o_i}^*$ does not exist, then the time to order is determined by the sign of the partial derivative $\partial E\left[TC_{ci}\right]/\partial t_{o_i}$. If it is positive, then the customer should order at time ($t_r - L_i$), and if it is negative, then the customer should order at $t = 0$.

4.4. Modeling and Solving the Supplier's Problem

In the two-echelon, single-supplier, *n*-customer problem with lateral transshipments, the role of the supplier does not change appreciably compared to that with the problem with no LTs. The supplier's long-run expected cost per unit time is computed based on the same expected number of parts that need to be produced for the entire system. In other words, the expected demand of the system is the same. The same situation applies to the expected number of emergency shipments for the system. The supplier's long-run expected total cost function consists of production, inventory holding, and emergency production/shipping costs divided by the expected cycle length, *i.e.*,

$$E\left[TC_s\right]=\frac{\text{Production Cost}+E[\text{Inventory Holding Cost}]\ +\ E[\text{Emergency Shipping Cost}]}{E[\text{Cycle Length}]}. \tag{4.16}$$

The supplier's cost structure is similar in both cases, with and without LTs. Production cost is incurred in every cycle for each customer, holding cost is incurred any time a part for a customer has been produced until the order from the customer arrives, and the emergency shipment cost at the supplier is incurred anytime a shipment is required when a part is not on-hand and ready to be shipped when it is ordered by a customer.

However, for the optimization problem, the shipping might vary according to which source of replenishment is selected by the customers. Although the supplier will not ship parts to customers using LTs, it has to produce those parts to replenish the inventory of the supplying customer. Therefore, more than one cost function is needed to account for each customer in the system. In other words, the long-run expected total cost of the supplier is calculated based on two options: (1) based on the time that the customer orders directly to the supplier t_o and the supplier produces to that customer t_p, and (2) based on the time that a customer orders an LT t_{o_i} and the time to produce to the supplying customer t_p.

The supplier's expected cost is expressed as

$$E[\text{PHE Cost}] = c_p + c_h^s\left(g\left(n\right)\right)\left(\int_{t_p+p}^{t_o} (t_o - t) f\left(t\right)\, dt + \left(t_o - p - t_p\right)\int_{t_o}^{\infty} f\left(t\right) dt\right) + c_e\left(1 - g\left(n\right)\right)\left(\int_{t_o}^{t_o+p} p\, f\left(t\right) dt\right) \tag{4.17}$$

$$E[\text{PHE Cost}_i] = c_p + c_h^s\left(g\left(n\right)\right)\left(\int_{t_p+p}^{t_{o_i}} (t_{o_i} - t) f\left(t\right) dt + \left(t_{o_i} - p - t_p\right)\int_{t_{o_i}}^{\infty} f\left(t\right) dt\right) + c_e\left(1 - g\left(n\right)\right)\left(\int_{t_{o_i}}^{t_{o_i}+p} pf\left(t\right) dt\right) \tag{4.18}$$

Since the supplier wants to minimize the long-run expected total cost, then the cycle length for the supplier has the two forms, that is,

$$E\left[\text{Cycle Length}\right] = g\left(n\right)\left(\int_0^{t_o} tf\left(t\right) dt + \int_{t_o}^{\infty} t_o f\left(t\right) dt\right) + \left(1 - g\left(n\right)\right)\left(\int_0^{t_p} \left(t + p\right) f\left(t\right) dt + \int_{t_p}^{t_o} \left(t_p + p\right) f\left(t\right) dt + \int_{t_o}^{\infty} \left(t_o + p\right) f\left(t\right) dt\right) \tag{4.19}$$

and

$$E\left[\text{Cycle Length}_i\right] = g\left(n\right)\left(\int_0^{t_{o_i}} tf\left(t\right) dt + \int_{t_{o_i}}^{\infty} t_{o_i} f\left(t\right) dt\right) + \left(1 - g\left(n\right)\right)\left(\int_0^{t_p} \left(t + p\right) f\left(t\right) dt + \int_{t_p}^{t_{o_i}} \left(t_p + p\right) f\left(t\right) dt + \int_{t_{o_i}}^{\infty} \left(t_{o_i} + p\right) f\left(t\right) dt\right). \tag{4.20}$$

Now, combining Eqs. 4.17 and 4.19 and combining Eqs. 4.18 and 4.20, the supplier's long-run expected total cost per unit time are

$$E[TC_s]=\frac{\left[c_p+c_h^s\big(g(n)\big)\left(\int_{t_p+p}^{t_o}(t_o-t)f(t)\,dt+\big(t_o-p-t_p\big)\int_{t_o}^{\infty}f(t)\,dt\right)+c_e\big(1-g(n)\big)\left(\int_{t_o}^{t_o+p}p\,f(t)\,dt\right)\right]}{\left[\begin{array}{l} g(n)\left(\int_0^{t_o}(t)\;f(t)\,dt\;+\int_{t_o}^{\infty}(t_o)\;f(t)\,dt\right)+\\ \big(1-g(n)\big)\left(\int_0^{t_p}(t+p)\;f(t)\,dt\;+\int_{t_p}^{t_o}\big(t_p+p\big)\;f(t)\,dt\;+\int_{t_o}^{\infty}(t_o+p)\;f(t)\,dt\right)\end{array}\right]} \tag{4.21}$$

and

$$E[TC_{s_i}]=\frac{\left[c_p+c_h^s\big(g(n)\big)\left(\int_{t_p+p}^{t_{o_i}}(t_{o_i}-t)f(t)\,dt+\big(t_{o_i}-p-t_p\big)\int_{t_{o_i}}^{\infty}f(t)\,dt\right)+c_e\big(1-g(n)\big)\left(\int_{t_{o_i}}^{t_{o_i}+p}p\,f(t)\,dt\right)\right]}{\left[\begin{array}{l} g(n)\left(\int_0^{t_{o_i}}(t)\;f(t)\,dt\;+\int_{t_o}^{\infty}(t_{o_i})\;f(t)\,dt\right)+\\ \big(1-g(n)\big)\left(\int_0^{t_p}(t+p)\;f(t)\,dt\;+\int_{t_p}^{t_{o_i}}\big(t_p+p\big)\;f(t)\,dt\;+\int_{t_{o_i}}^{\infty}(t_{o_i}+p)\;f(t)\,dt\right)\end{array}\right]}. \tag{4.22}$$

<u>Derivation of the Supplier Inventory and Production Policies</u>

Only one of the supplier's total cost functions will be used (Eq. 4.21) to show the results of the derivation since they have the same structure, the only difference is the time to order variable (t_o and t_{o_i}). Then, the long-run expected total cost per unit time at the supplier, assuming an exponential part failure distribution, is

$$E[TC_s]=\frac{\left[\begin{array}{l} e^{\lambda\,(t_o+t_p)}\lambda\Big(c_p-c_e e^{-\lambda\,(t_o+p)}p\big(e^{\lambda p}-1\big)\big(g(n)-1\big)\Big)+\\ \left(\frac{1}{\lambda}\right)\Big(c_h^s\;g(n)e^{-\lambda\,(t_o+t_p+p)}\Big(e^{\lambda t_o}\big(-\lambda\big(p-t_o+t_p\big)-1\big)+e^{\lambda\,(t_p+p)}\big(1-\lambda\big(p-t_o+t_p\big)\big)\Big)\Big)\end{array}\right]}{e^{\lambda\,t_o}\big(g(n)-1\big)+e^{\lambda\,t_p}\Big(\lambda\big(t_o-t_p\big)+g(n)\big(t_p\lambda-t_o\lambda-1\big)\Big)+e^{\lambda\,(t_o+t_p)}\Big(1+p\big(\lambda-g(n)\big)\lambda\Big)}. \tag{4.23}$$

Once more, the optimal values of t_o, t_{o_i} and t_p are found by setting the partial derivatives $\partial E\left[TC_s\right]/\partial t_o$ and $\partial E\left[TC_s\right]/\partial t_p$ equal to zero and solving the resulting equations simultaneously yielding

$$t_o^* = \frac{\ln\left[\dfrac{e^{-\lambda p}\left(c_h^s\, g(n)\, e^{2\lambda p} + c_e\, p\, \lambda\left(e^{\lambda p}-1\right)^2\left(g(n)-1\right) - c_h^s\left(g(n)-1\right)\left(1+p\lambda\right)\right)}{c_h^s + c_p\, \lambda\, \left(e^{\lambda p}-1\right) + c_h^s\, p\, \lambda}\right]}{\lambda} \tag{4.24}$$

and

$$t_p^* = \frac{\ln\left[\dfrac{g(n)\left(c_h^s\, g(n)\, e^{2\lambda p} + c_e\, p\, \lambda\left(e^{\lambda p}-1\right)^2\left(g(n)-1\right) - c_h^s\left(g(n)-1\right)\left(1+p\lambda\right)\right)}{\begin{array}{l} e^{\lambda p}\, \lambda\, \left(g(n)-1\right)\left(c_p - p\, c_e\right)\left(1+p\,\lambda\right) + \\ e^{2\lambda p}\left(c_h^s\, \left(g(n)+g(n)\, p\, \lambda\right) + \lambda\left(c_p\, g(n) + c_e\, p\left(g(n)-1\right)\left(1+p\,\lambda\right)\right)\right) \end{array}}\right]}{\lambda}. \tag{4.25}$$

The same insights obtained in the customer's problem apply to the supplier's expected total cost function. The determinant of the Hessian matrix is complex, and joint convexity over the two variables cannot be shown. However, the expected total cost function for the supplier is pseudoconvex in t_o. Consequently, for a given t_p, the optimal t_o can either be $(t_p + p)$, ∞ or t_o^*. In case t_o^* does not exist, the partial derivative $\partial E\left[TC_s\right]/\partial t_o$ will determine the best time to receive an order. If this derivative is negative, then the best ordering time should be at time zero. If it is positive, the best time to receive an order is at $(t_p + p)$. The expected total cost function is also pseudoconvex in t_p. So, for a given t_o, which is typically the case, the optimal t_p can either be 0, $(t_o - p)$ or t_p^*. If t_p^* does not exist, then the time to produce a part is determined by the sign of the partial derivative $\partial E\left[TC_s\right]/\partial t_p$. If it is positive, the supplier should produce a part at $(t_o - p)$. On the other hand, if it is negative, then the supplier should produce a part at $t = 0$.

We describe the joint inventory-maintenance optimization problem within the two-echelon supply network and lateral transshipments, where the individual long-run expected total cost per unit time to the customer-level problem and the supplier-level problem are derived separately. In the next section, the results of the simultaneous optimization is presented and compared with those from the supply network configuration with no LTs.

4.5. Computational Study

This section shows the performance of the multiobjective optimization approach for joint inventory and maintenance policy optimization with lateral transshipments. Under this approach the optimal values of t_o^*, $t_{o_i}^*$, t_r^* and t_p^* that minimize the total system-wide cost are identified to determine the best ordering policy for the overall system. We develop experiments for the case with a two-customer configuration (n = 2). $E[TC_s]$, $E\left[TC_{s_i}\right]$, $E[TC_c]$ and $E\left[TC_{c_i}\right]$ are the objective functions used by the NSGA-II (Deb *et al.*, 2002) for the simultaneous optimization.

Table 4.1 summarizes the full factorial experimental design of the problem instances. The values and ranges of the parameters are the same values used in the experiments for the model without LTs. For the single service part, a failure distribution of exponential form with a rate parameter $\lambda = 0.01$, or a mean time between part failures $\mu = 100$, is used.

Table 4.1. Ranges of the problem instance parameters for a two-echelon single-supplier two-customer service parts inventory supply chain system.

	Supplier	Customer 1	Customer 2
Unit Ordering Cost, c_o	-	10	15
Unit Ordering Cost, c_{o_i}		5	5
Unit Holding Cost, c_h	[300, 600]	[300, 600]	[300, 600]
Unit Replacement Cost, c_r	-	30	25
Unit Failure Cost, c_f	-	28	20
Unit Shortage Cost, c_s	-	[300, 600]	[300, 600]
Mean Time Between Failure, μ	-	100	100
Unit Emergency Shipping Cost, c_e	[300, 600]	-	-
Unit Production Cost, c_p	50	-	-
Unit Production Time, p	4	-	-
Order Delivery Lead Time, L	-	2	3
Order Delivery Lead Time, L_i		1	1

The search control parameters for the real-coded NSGA-II are summarized in Table 4.2. These parameters are the similar to those used in the computational study where no LTs are allowed, so that there is common basis for comparison of the supply chain configurations. The specific parameter values used in the LT model configuration are summarized in Table 4.3. The solution with the lowest expected total cost per unit time and the associated t_o or t_{o_2} , t_r and t_p values are reported. Table 4.4 summarizes the replacement and ordering policies for the customers and the production schedule for the supplier. In the case that t_{o_1} or t_{o_2} generates the lowest total cost, t_o is not taken into consideration, and vice versa.

Table 4.2. Search control parameters for NSGA-II.

Parameter	Value
Population Size, P	200
Number of Generations, G	10,000
Crossover Rate, p_c	1.000
Mutation Rate, p_m = 1 / number of vars)	0.167
Distribution Index for Crossover η_c	10
Distribution Index for Mutation η_m	25

Table 4.3. Specific problem instance cost parameters for a two-echelon, one supplier two customer service parts inventory supply chain system with lateral transshipments.

Prob Inst	c_p S	c_h S	c_e S	c_r C1	c_r C2	c_f C1	c_f C2	c_o C1	c_o C2	c_{o_1} C2	c_{o_2} C1	c_h C1	c_h C2	c_s C1	c_s C2
1	50	300	300	30	25	28	20	10	15	5	5	300	300	300	300
2	50	600	300	30	25	28	20	10	15	5	5	300	300	300	300
3	50	300	300	30	25	28	20	10	15	5	5	600	300	300	300
4	50	600	300	30	25	28	20	10	15	5	5	600	300	300	300
5	50	300	300	30	25	28	20	10	15	5	5	300	600	300	300
6	50	600	300	30	25	28	20	10	15	5	5	300	600	300	300
7	50	300	300	30	25	28	20	10	15	5	5	600	600	300	300
8	50	600	300	30	25	28	20	10	15	5	5	600	600	300	300
9	50	300	300	30	25	28	20	10	15	5	5	300	300	600	300
10	50	600	300	30	25	28	20	10	15	5	5	300	300	600	300
11	50	300	300	30	25	28	20	10	15	5	5	600	300	600	300
12	50	600	300	30	25	28	20	10	15	5	5	600	300	600	300
13	50	300	300	30	25	28	20	10	15	5	5	300	600	600	300
14	50	600	300	30	25	28	20	10	15	5	5	300	600	600	300
15	50	300	300	30	25	28	20	10	15	5	5	600	600	600	300
16	50	600	300	30	25	28	20	10	15	5	5	600	600	600	300
17	50	300	300	30	25	28	20	10	15	5	5	300	300	300	600
18	50	600	300	30	25	28	20	10	15	5	5	300	300	300	600
19	50	300	300	30	25	28	20	10	15	5	5	600	300	300	600
20	50	600	300	30	25	28	20	10	15	5	5	600	300	300	600
21	50	300	300	30	25	28	20	10	15	5	5	300	600	300	600
22	50	600	300	30	25	28	20	10	15	5	5	300	600	300	600
23	50	300	300	30	25	28	20	10	15	5	5	600	600	300	600
24	50	600	300	30	25	28	20	10	15	5	5	600	600	300	600
25	50	300	300	30	25	28	20	10	15	5	5	300	300	600	600
26	50	600	300	30	25	28	20	10	15	5	5	300	300	600	600
27	50	300	300	30	25	28	20	10	15	5	5	600	300	600	600
28	50	600	300	30	25	28	20	10	15	5	5	600	300	600	600
29	50	300	300	30	25	28	20	10	15	5	5	300	600	600	600
30	50	600	300	30	25	28	20	10	15	5	5	300	600	600	600
31	50	300	300	30	25	28	20	10	15	5	5	600	600	600	600
32	50	600	300	30	25	28	20	10	15	5	5	600	600	600	600
33	50	300	600	30	25	28	20	10	15	5	5	300	300	300	300
34	50	600	600	30	25	28	20	10	15	5	5	300	300	300	300
35	50	300	600	30	25	28	20	10	15	5	5	600	300	300	300
36	50	600	600	30	25	28	20	10	15	5	5	600	300	300	300
37	50	300	600	30	25	28	20	10	15	5	5	300	600	300	300
38	50	600	600	30	25	28	20	10	15	5	5	300	600	300	300
39	50	300	600	30	25	28	20	10	15	5	5	600	600	300	300
40	50	600	600	30	25	28	20	10	15	5	5	600	600	300	300
41	50	300	600	30	25	28	20	10	15	5	5	300	300	600	300
42	50	600	600	30	25	28	20	10	15	5	5	300	300	600	300

Table 4.3. (cont'd) Specific problem instance cost parameters for a two-echelon, one supplier two customer service parts inventory supply chain system with lateral transshipments.

Prob Inst	c_p S	c_h S	c_e S	c_r C1	c_r C2	c_f C1	c_f C2	c_o C1	c_o C2	c_{o_1} C2	c_{o_2} C1	c_h C1	c_h C2	c_s C1	c_s C2
43	50	300	600	30	25	28	20	10	15	5	5	600	300	600	300
44	50	600	600	30	25	28	20	10	15	5	5	600	300	600	300
45	50	300	600	30	25	28	20	10	15	5	5	300	600	600	300
46	50	600	600	30	25	28	20	10	15	5	5	300	600	600	300
47	50	300	600	30	25	28	20	10	15	5	5	600	600	600	300
48	50	600	600	30	25	28	20	10	15	5	5	600	600	600	300
49	50	300	600	30	25	28	20	10	15	5	5	300	300	300	600
50	50	600	600	30	25	28	20	10	15	5	5	300	300	300	600
51	50	300	600	30	25	28	20	10	15	5	5	600	300	300	600
52	50	600	600	30	25	28	20	10	15	5	5	600	300	300	600
53	50	300	600	30	25	28	20	10	15	5	5	300	600	300	600
54	50	600	600	30	25	28	20	10	15	5	5	300	600	300	600
55	50	300	600	30	25	28	20	10	15	5	5	600	600	300	600
56	50	600	600	30	25	28	20	10	15	5	5	600	600	300	600
57	50	300	600	30	25	28	20	10	15	5	5	300	300	600	600
58	50	600	600	30	25	28	20	10	15	5	5	300	300	600	600
59	50	300	600	30	25	28	20	10	15	5	5	600	300	600	600
60	50	600	600	30	25	28	20	10	15	5	5	600	300	600	600
61	50	300	600	30	25	28	20	10	15	5	5	300	600	600	600
62	50	600	600	30	25	28	20	10	15	5	5	300	600	600	600
63	50	300	600	30	25	28	20	10	15	5	5	600	600	600	600
64	50	600	600	30	25	28	20	10	15	5	5	600	600	600	600

Table 4.4. Service part ordering, part replacement, part production times and total costs when the cost objectives are optimized simultaneously.

Prob Inst	t_o C1	t_{o_2} C1	t_r C1	t_p C1	t_o C2	t_{o_1} C2	t_r C2	t_p C2	Cost/Time S	Cost/Time C1	Cost/Time C2	Long-Run Total Cost/Time
1	4543	-	4987.9	4336.9	-	4586.8	4904.79	4381.2	0.99	8.10	3.47	12.56
2	4547	-	4562.8	4486.3	-	4227.5	4959.69	4213.7	0.99	8.10	3.47	12.56
3	4378	-	4910.7	4327	-	4013.5	4825.48	4007.7	0.99	33.04	13.37	47.40
4	4414.3	-	4701.2	4406.8	-	3685.8	3694.77	3681.8	0.99	33.04	13.37	47.40
5	4209	-	4970.8	4183.9	-	4474.4	4904.52	4336.2	1.00	7.06	3.47	11.52
6	-	4157.2	4694	4152.7	4177.8	-	4231.5	4171.6	1.00	3.59	9.77	14.36
7	4411.4	-	4910.2	4391.4	-	4589.3	4978.25	4566.1	1.00	7.06	3.47	11.52
8	4443.8	-	4551.6	4414.6	-	4602.1	4898.97	4457.8	1.00	7.06	3.47	11.52
9	-	4642	4845.2	4577.5	4116.4	-	4183.6	4107.3	1.00	6.56	9.77	17.33
10	-	4083.8	4983.2	4079.4	683.77	-	688.03	679.05	1.02	6.56	9.77	17.35
11	-	4710.2	4922.5	4539.7	1649.2	-	1653.93	1494	1.00	6.56	9.77	17.33
12	-	4157	4950.6	4149.9	4649.8	-	4773.76	4528	1.00	6.56	9.77	17.33
13	-	4509.6	4956.8	4378.8	4123.2	-	4170.64	4112.7	1.00	6.56	9.77	17.33
14	1113.8	-	1117.9	1090.9	-	4373.6	4860.89	4335.7	1.00	13.46	3.47	17.93
15	-	4151.8	4874.4	4132.4	4672.5	-	4778.18	4641.3	1.00	6.56	9.77	17.33
16	-	4226.1	4974.5	4167.9	973.52	-	977.822	967.67	1.00	6.56	9.77	17.33
17	4122.1	-	4507.7	4065.3	-	4411.9	4870.37	4360.2	1.00	7.06	6.44	14.49
18	4368.5	-	4658.4	4345.6	-	4558.6	4798.88	4497.1	1.00	7.06	6.44	14.49
19	4648.8	-	4915.4	4569.8	-	4303.3	4821.72	4285.3	1.00	7.06	6.44	14.49
20	4330.1	-	4656.9	4266.8	-	4342.7	4866.35	4325.2	1.00	7.06	6.44	14.49
21	4269.2	-	4620.2	4258.4	-	4399.6	4932.07	4362.5	1.00	7.06	6.44	14.49
22	4214.3	-	4303.1	4201.4	-	4431.2	4977.6	4357.7	1.00	7.06	6.44	14.49
23	4494.4	-	4785.5	4378	-	4254.8	4731.6	4221.8	1.00	7.06	6.44	14.49
24	4615.3	-	4667.6	4363.9	-	4234.7	4897	4222.6	1.00	7.06	6.44	14.49
25	896.42	-	899.8	883.87	-	4503	4916.77	4356.9	1.00	13.47	6.44	20.90
26	4456.6	-	4490.3	4354.1	-	4641	4686.39	4471	1.00	13.46	6.44	20.90
27	4100.9	-	4123.9	4093.2	-	4484.9	4955.02	4328.9	1.00	13.46	6.44	20.90
28	4492.4	-	4655	4368.8	-	4121.6	4946.61	4115.7	1.00	13.46	6.44	20.90
29	4784	-	4939.5	4510.3	-	4252.6	4849.44	4225.8	1.00	13.46	6.44	20.90
30	4291.6	-	4473	4273.1	-	4200.5	4605.99	4193.1	1.00	13.46	6.44	20.90
31	4388.2	-	4440.8	4380.7	-	4188.9	4966.96	4177.3	1.00	13.46	6.44	20.90
32	2148.4	-	2159	2142.3	-	4510.1	4951.1	4468.7	1.00	13.46	6.44	20.90
33	4802.4	-	4851.3	4436.6	-	4459.3	4755.63	4371	1.00	7.06	3.47	11.52

Table 4.4. (Cont'd) Service part ordering, part replacement, part production times and total costs when the cost objectives are optimized simultaneously.

Prob Inst	t_o C1	t_{o_2} C1	t_r C1	t_p C1	t_o C2	t_{o_1} C2	t_r C2	t_p C2	Cost/Time S	Cost/Time C1	Cost/Time C2	Long-Run Total Cost/Time
34	4115.2	-	4645.3	4107.6	-	4736.3	4989.85	4552.4	1.00	7.06	3.47	11.52
35	4461.8	-	4693.8	4353.8	-	4614.5	4718.66	4452.1	1.00	7.06	3.47	11.52
36	4232.9	-	4442.2	4210.8	-	4550.1	4951.83	4401.5	1.00	7.06	3.47	11.52
37	4160.9	-	4900.5	4141.4	-	4395.2	4943.71	4339.8	1.00	7.06	3.47	11.52
38	4301.6	-	4503.7	4271.7	-	4732.9	4882.22	4455.2	1.00	7.06	3.47	11.52
39	4198.8	-	4231.2	4143.2	-	4412.4	4962.29	4321.6	1.00	7.06	3.47	11.52
40	4617.3	-	4907.6	4611.8	-	4271.2	4760.45	4260.4	1.00	7.06	3.47	11.52
41	-	4738.6	4788.1	4441.3	4578.9	-	4771.09	4490.2	1.00	6.56	9.77	17.33
42	-	4147.7	4908	4140.5	4814.8	-	4908.09	4329.9	1.00	6.56	9.77	17.33
43	-	4705.2	4821.2	4574.7	4441.1	-	4651.48	4341.3	1.00	6.56	9.77	17.33
44	-	4175.5	4913.6	4156.6	4981.4	-	4991.66	4809.1	1.00	6.56	9.77	17.33
45	-	4529.5	4800.3	4345.5	4584	-	4751.56	4518.2	1.00	6.56	9.77	17.33
46	-	4082.9	4908	4078	4579.4	-	4980.56	4354.7	1.00	6.56	9.77	17.33
47	-	4614.9	4848.1	4562.7	4482.8	-	4782.58	4397.6	1.00	6.56	9.77	17.33
48	-	4159.6	4894.8	4153.2	4506	-	4913.8	4491	1.00	6.56	9.77	17.33
49	4574.8	-	4777.3	4434.1	-	4546.2	4761.43	4386.3	1.00	7.06	6.44	14.49
50	4404	-	4900.6	4293.3	-	4426.8	4931.31	4401.8	1.00	7.06	6.44	14.49
51	1969	-	1972.3	1963	-	4689.6	4885.86	4342.1	1.00	7.06	6.44	14.49
52	1614.7	-	1618	1610.4	-	4087.7	4536.2	4083.3	1.00	7.06	6.44	14.49
53	4580.4	-	4667.4	4389.8	-	4493	4917.52	4342.5	1.00	7.06	6.44	14.49
54	4129.5	-	4938.3	4105.7	-	4577.2	4739.22	4454.5	1.00	7.06	6.44	14.49
55	2063.5	-	2067	2057.4	-	4150.9	4945.24	4140.1	1.00	7.06	6.44	14.49
56	4345	-	4613.2	4258.9	-	4165.3	4845.53	4158.3	1.00	7.06	6.44	14.49
57	4603.2	-	4881.1	4307.7	-	4185.4	4864.41	4175.2	1.00	13.46	6.44	20.90
58	658	-	661.55	651.36	-	4620.4	4972.19	4465.1	1.04	13.47	6.44	20.94
59	4682.9	-	4800.9	4513.7	-	4382.9	4814.12	4293.7	1.00	13.46	6.44	20.90
60	4066.9	-	4072.8	4059.2	-	4577	4884.98	4464.6	1.00	13.46	6.44	20.90
61	4462.4	-	4624.6	4357.4	-	4305	4706.95	4209.9	1.00	13.46	6.44	20.90
62	4779.8	-	4849.6	4490.8	-	4578.8	4855.4	4492.6	1.00	13.46	6.44	20.90
63	4852.2	-	4942.4	4500.6	-	4476.9	4869.33	4382.5	1.00	13.46	6.44	20.90
64	4639.3	-	4720.2	4537.9	-	4529.3	4839.79	4490.5	1.00	13.46	6.44	20.90

General Behavior of the Set of Pareto Optima

Figure 4.2 shows the Pareto efficient frontier for Customer 1 vs. Customer 2 for Prob Inst 44. Figure 4.3 shows the Pareto efficient frontier for Customer 1 vs. Customer 2 for Prob Inst 52. Figure 4.4 shows the Customer 1 vs. Customer 2 vs. the Supplier. It can be seen that there is a reasonable level of diversity among the solutions.

From Table 4.4, we can see that in all the cases using LTs is more economically beneficial than not using them. Table 4.5 compares the performance of the supply chain model when using and not using lateral transshipments. The average improvement over the non-LT model over the 64 problem instances is 36.79%, with a standard deviation of 6.81%, and ranging from 19.44% to 46.36%.

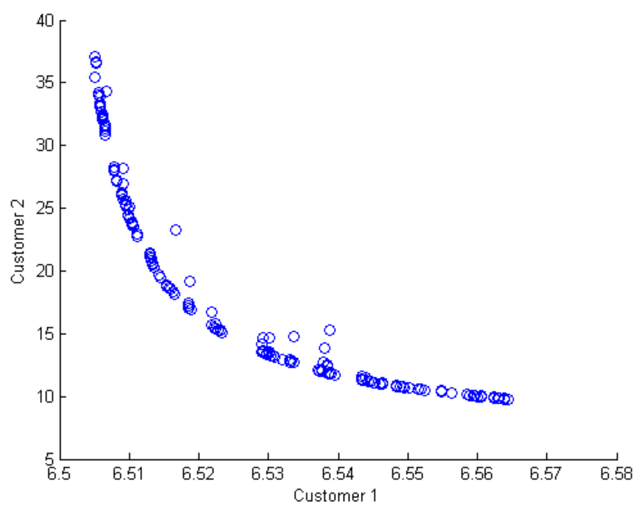

Figure 4.2. Customer 1 long-run expected total cost per unit time vs. Customer 2 long-run expected total cost per unit time for Prob Inst 44.

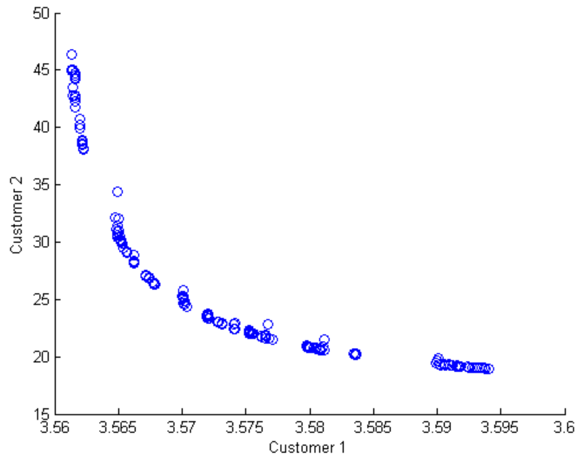

Figure 4.3. Customer 1 long-run expected total cost per unit time vs. Customer 2 long-run expected total cost per unit time for Prob Inst 52.

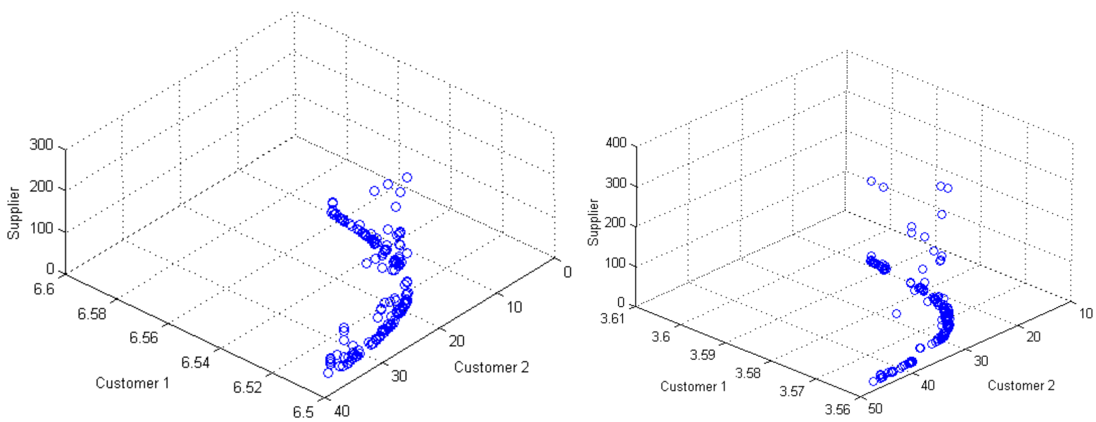

Figure 4.4. Customer 1 long-run expected total cost per unit time vs. Customer 2 long-run expected total cost per unit time vs. Supplier long long-run expected total cost per unit time for Prob Insts 44 and 52.

Table 4.5.Comparison of the models with and without lateral transshipments.

Prob Inst	Long-Run Total Cost per Unit Time with LTs allowed	Long-Run Total Cost per Unit Time with No LTs allowed	Difference	% Improve
1	12.56	19.78	7.21	36.47%
2	12.56	19.78	7.21	36.47%
3	47.40	78.49	31.09	39.61%
4	47.40	78.49	31.09	39.61%
5	11.52	17.82	6.30	35.35%
6	14.36	17.82	3.46	19.44%
7	11.52	17.82	6.30	35.35%
8	11.52	17.82	6.30	35.35%
9	17.33	24.23	6.90	28.48%
10	17.35	24.23	6.88	28.39%
11	17.33	24.23	6.90	28.48%
12	17.33	24.23	6.90	28.48%
13	17.33	24.23	6.90	28.48%
14	17.93	24.23	6.30	25.99%
15	17.33	24.23	6.90	28.48%
16	17.33	24.23	6.90	28.47%
17	14.492	27.02	12.52	46.36%
18	14.492	27.02	12.52	46.36%
19	14.492	27.02	12.52	46.36%
20	14.492	27.02	12.52	46.36%
21	14.492	27.02	12.52	46.36%
22	14.492	27.02	12.52	46.36%
23	14.492	27.02	12.52	46.36%
24	14.492	27.02	12.52	46.36%
25	20.90	33.42	12.52	37.46%
26	20.90	33.42	12.52	37.47%
27	20.90	33.42	12.52	37.47%
28	20.90	33.42	12.52	37.47%
29	20.90	33.42	12.52	37.47%
30	20.90	33.42	12.52	37.47%
31	20.90	33.42	12.52	37.47%
32	20.90	33.42	12.52	37.47%
33	11.52	17.82	6.30	35.35%
34	11.52	17.82	6.30	35.35%
35	11.52	17.82	6.30	35.35%
36	11.52	17.82	6.30	35.35%
37	11.52	17.82	6.30	35.35%
38	11.52	17.82	6.30	35.35%
39	11.52	17.82	6.30	35.35%
40	11.52	17.82	6.30	35.35%
41	17.33	24.23	6.90	28.48%
42	17.33	24.23	6.90	28.48%

Table 4.5. (cont'd) Comparison of the models with and without lateral transshipments.

Prob Inst	Long-Run Total Cost per Unit Time with LTs allowed	Long-Run Total Cost per Unit Time with No LTs allowed	Difference	% Improve
43	17.33	24.23	6.90	28.48%
44	17.33	24.23	6.90	28.48%
45	17.33	24.23	6.90	28.48%
46	17.33	24.23	6.90	28.48%
47	17.33	24.23	6.90	28.48%
48	17.33	24.23	6.90	28.48%
49	14.49	27.02	12.52	46.36%
50	14.49	27.02	12.52	46.36%
51	14.49	27.02	12.52	46.36%
52	14.49	27.02	12.52	46.36%
53	14.49	27.02	12.52	46.36%
54	14.49	27.02	12.52	46.36%
55	14.49	27.02	12.52	46.36%
56	14.49	27.02	12.52	46.36%
57	20.90	33.42	12.52	37.47%
58	20.94	33.42	12.48	37.34%
59	20.90	33.42	12.52	37.47%
60	20.90	33.42	12.52	37.47%
61	20.90	33.42	12.52	37.47%
62	20.90	33.42	12.52	37.47%
63	20.90	33.42	12.52	37.47%
64	20.90	33.42	12.52	37.47%

4.6. Summary and Conclusions

In this chapter, we develop cost models for a joint maintenance and service parts supply chain system in a two-echelon, single supplier, and *n*-customers configuration where the customers share their inventories and lateral transshipments are allowed. We use a multiobjective optimization approach to determine the ordering and replacement policies that minimize the system-wide long-run expected total cost per unit time.

In a computational study, we use a full factorial experimental design and create 64 different problem instances to evaluate the performance of supply chain system. We also compare the supply chain system that allows LTs with a supply chain system that does not allow lateral transshipments between customers. The results of the comparison reveal important differences that demonstrate the benefits of the use of lateral transshipments and the joint optimization among all members of the supply chain.

One of the most important conclusions about the joint optimization approach is that decision-makers are presented with a variety of alternatives that can be considered with reasonable benefit for the whole system. These solutions can be negotiated with the stakeholders so that all of them can be included in the process and better relationships can be established.

CHAPTER 5:
SUMMARY OF RESEARCH AND FUTURE RESEARCH DIRECTIONS

5.1. Summary of This Research Investigation

Maintenance and service parts inventories are two of the three important areas of this investigation. We develop mathematical cost models that represent the long-run expected total cost per unit time for the stakeholders involved in a multi-echelon service parts supply chain. These formulations are addressed from two different perspectives – the supplier's and the customer's. From the supplier's point of view, the most important parameters relative to inventory management and production are included in the model formulation. From the customer's point of view, the most important parameters for service parts inventory management and maintenance are considered.

Traditionally, in a multi-echelon service parts supply chain, all the stakeholders act as separate entities seeking to minimize their individual costs without considering other customers or the supplier. This is considered the traditional, or decoupled, approach in this research. In an effort to improve the performance of the service parts supply chain, we present a multiobjective modeling approach in which all the stakeholders' costs are minimized simultaneously. For the decoupled and the proposed optimization approaches, we employ evolutionary algorithm procedures. In the case of the decoupled approach, a real-coded genetic algorithm is used. Alternatively, for the proposed multiobjective approach, NSGA-II is used.

These two approaches are compared in different supply chain network configurations where the number of customers in the supply chain is varied. For the first experiment, the configuration involves only one customer and a parts supplier. We conclude that the simultaneous optimization of the objectives yields results superior to that of the decoupled approach. The second experiment involves one supplier and two customers in the network. The results of this configuration are similar to those of the single-supplier, single-customer configuration. In other words, simultaneous optimization results in greater economic benefit to the service parts supply chain.

We extend the previous model formulation by relaxing one of the more important assumptions. For this new situation, we present a new formulation that reflects the inclusion of lateral transshipments. Since it is demonstrated that simultaneous optimization is the approach to use in this type of optimization problem, we employ the same optimization procedures for this proposed model formulation. We use the results from the previous model with two customers in the system as a benchmark to evaluate the use of lateral transshipments in the supply chain.

We show, like many other researchers, that allowing shipments between customers yields significant improvements to a service parts inventory supply chain. It is clear that those improvements primarily come from the reduced shortage costs because wider areas can be covered with inventory pooling. In other words, the response time, when an unexpected failure occurs, is reduced. In addition, there is also a significant reduction in the ordering cost and holding costs. This is due to the proximity of the source of replenishment and also the average of replenishments that have to be performed.

5.2. Directions for Future Research

There are several additional aspects that can be considered and investigated in order to improve these models and make them more applicable to different real-world situation. Thus, extensions to this research are now presented.

5.2.1. Additional and Improved Stochastic Parameters

Some parameters in the model are assumed to be deterministic. Nonetheless, they can be associated with a probability distribution. In the service parts inventory world, there are several assumptions about these parameters that are valid for some cases, but in reality, they do not work that way. For instance, a deterministic lead time is often assumed. However, it is clear that in the real-world, a deterministic lead time almost never occurs. Another example is the unit production time that falls into the same category. Thus, a probability distribution that represents the behavior for these parameters should be taken into account, and included in the models described in Eqs. 3.6, 3.16, 4.10, 4.11 and 4.22. For a real-world problem, real data need to be collected and a better representation of the parameters used for the determination of the long-run expected cost functions.

5.2.2. Statistical Analysis of the Decision Variables

An evolutionary algorithm random procedure is used for the multiobjective optimization in this research. Thus, the results of a single optimization run are not always going to be the same. For this reason, it is necessary to develop a statistical analysis methodology for the interpretation and use of the results given by this type of algorithms. A Pareto front is a set of non-dominated solutions that an evolutionary algorithm yields in one replication. To have a more valid set of solutions several replications need to be performed, so confidence intervals can be calculated for each of the solutions in the Pareto front. It is at this point where a methodology for the study of these results is needed, and better conclusions can be inferred.

5.2.3. Extend to Several Service Parts

The models derived in this research are limited to a single part. This assumption is made because the application of the model is intended for very expensive service parts. For instance, in the power generation industry, turbines could cost millions of dollars. One possible extension of the model is to include situations where more than one part could be ordered and held in stock. This type of configuration could be very useful in the military or aviation industry where several expensive parts are managed.

5.2.4. Improve the Performance for Large Number of Objectives

In order to make this investigation more applicable to real-world problems, the accuracy of the results needs to be tested against the number of objectives in the system. The effect of an increasing number of objectives should be evaluated to guarantee that the inclusion of them does not affect the precision of the results. In other words, when having a large number of objectives, find a methodology to evaluate how the results are deviated from reality. For this, performance metrics need to be investigated that analyze the number of objectives that are optimized simultaneously. Therefore, a study of the multiobjective algorithms used for service parts inventory optimization needs to be performed.

5.2.5. Develop a Decision Support Application

The ultimate goal of this dissertation is to set the basis for the decision-making process in a multi-echelon service parts supply chain. A very interesting and useful task is to develop an application that integrates all the features developed in this dissertation. First of all, since the models derived in this research are unique, an application where the objective functions can be adapted to different configurations would be an important contribution for the industry. In addition, the handling of the different parameters could be also an important feature. The parameters can be analyzed with integrated statistical tools to fit theoretical or empirical probability distributions that the model should be able to operate. Finally, this application should include multiobjective optimization and have a very dynamic and easy to understand representation of the results where sensitivity analysis can be performed.

LIST OF REFERENCES

Alfredsson, P. and Verrijdt, J. (1999). Modeling emergency supply flexibility in a two-echelon inventory system. *Management Science* **45**, 1416-1431.

Al-Rifai, M.H. and Rossetti, M.D. (2007). An efficient heuristic optimization algorithm for a two-echelon (*R*, *Q*) inventory system. *International Journal of Production Economics*, **109**, 195-213.

Archibald, T.W., Sassen, S.A. and Thomas, L.C. (1997). An optimal policy for a two depot inventory problem with stock transfer. *Management Science* **43**, 173-183.

April, J., Glover, F., Kelly, J. and Laguna, M. (2003). Practical introduction to simulation optimization. In: Chick, S., (Ed.). *Proceedings of the 2003 Winter Simulation Conference, Piscataway, NJ: Institute of Electrical and Electronics Engineers*, 71-78

Armstrong, M.J. and Arkins, D.R. (1996). Joint optimization of maintenance and inventory policies for a simple system. *IIE Transactions* **28**, 415-424.

Arnold, J. and Köchel, P. (1996). Evolutionary Optimization of a Multi-location Inventory Model with Lateral Transshipments. In *9th International Working Seminar on Production Economics* **2**, 401-412.

Avsar, Z.M. and Zijm, W.H.M. (2000). Resource constrained two-echelon inventory models for repairable item systems. Working paper, University of Twente, The Netherlands.

Axsäter, S. (1990a). Simple solution procedures for a class of two-echelon inventory problems. *Operations Research* **38**, 64-69.

Axsäter, S. (1990b). Modeling emergency lateral transshipments in inventory systems. *Management Science* **36**, 1329-1338.

Axsäter, S. (2003). Evaluation of unidirectional lateral transshipments and substitutions in inventory systems. *European Journal of Operational Research*, 149, 438-447.

Caglar, D., Li, C. and Simchi-Levi, D. (2004). Two-echelon spare parts inventory system subject to a service constraint. *IIE Transactions* **36**, 655-666.

Carnero, M.C. (2004) The control of the setting up of a predictive maintenance program using a system of indicators. *The International Journal of Management Science* **32**, 57-75.

Cassady, C.R., Murdock, W.P. and Polh, E.A. (2001). Selective maintenance for support equipment involving multiple maintenance actions. *European Journal of Operational Research* **129**, 252-258.

Coello, C.A.C., van Veldhuizen, D.A. and Lamont, G.B. (2002). Evolutionary Algorithms for Solving Multi-Objective Problems, (1st Ed.), New York: Kluwer Academic Publishers.

Cohen, M., Kamesam, P.V., Kleindorfer, P., Lee, H. and Tekerian, A. (1990). Optimizer: IBM's multiechelon inventory system for managing service logistics, *Interfaces* **20**, 65-82.

Das, T. and Sarkar, S. (1999). Optimal preventive maintenance in a production inventory system. *IIE Transactions* **31**, 537-551.

Deb, K., Pratap, A., Agarval, S. and Meyarivan, T.A. (2002). Fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* **6**, 182-197.

De Smidt-Destombes, K., Van der Heijden, M. and Van Harten A. (2005). On the interaction between maintenance, spare part inventories and repair capacity for a *k*-out-of-*N* system with wear-out. *European Journal of Operational Research*. (in press).

Diaz, A. (2003). Modeling approaches to optimize spares in multi-echelon systems. *International Journal of Logistics: Research and Applications* **6**, 51-62.

Diaz, A. (2005). Spare parts and maintenance optimization in a mobile telephone company. Working Paper.

Diaz, A. and Fu, M.C. (1997). Models for multi-echelon repairable item inventory systems with limited repair capacity", *European Journal of Operational Research* **97**, 480-492.

Diaz, A. and Fu, M.C. (2004). Multi-echelon models for repairable items: A review. Working Paper.

Feldman, R. M., & Valdez-Flores, C. (1995). *Applied Probability and Stochastic Processes* (1st ed., p. 320). Pws Pub Co.
Medhi, J. (1982). *Stochastic Processes.* (p. 400). Wiley Eastern 1982.

Goldberg, D.E. (1989). Genetic Algorithms in search, optimization, and machine learning. Addison- Wesley, New York

Eskandari, H. and Geiger, C.D. (2007). A fast Pareto genetic algorithm approach for solving expensive multiobjective optimization problems. *Journal of Heuristics* (in press).

United States General Accountability Office (GAO). (2003). Report on defense inventory.

Grahovac, J. and Chakravarty, A. (2001). Sharing and lateral transshipment of inventory in a supply chain with expensive low-demand items. *Management Science* **47**, 579-594.

Graves, S. A. (1985). Multi-echelon inventory model for a repairable item with one-for-one replenishment. *Management Science* **31**, 1247-1256.

Holland, J. (1975). *Adaptation in Nature and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.

Kennedy, W.J., Patterson, J.W., Fredendall, L.D. (2002). An overview of recent literature on spare parts inventories. *International Journal of production economics* **76**, 201-215.

Kilpi, J. and Vepsäläinen, A. (2004). Pooling of spare components between airlines. *Journal of Air Transport Management* **10**, 137-146.

Knowles, J. (2006). ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, **10(1)**, 50-66.

Knowles, J. and Corne, D. (1999). The Pareto archived evolution strategy: A new baseline algorithm for multiobjective optimization. *Proceedings of the 1999 Congress on Evolutionary Computation, Washington, D.C., July 1999, IEEE Service Center*, 98-105.

Koçağa, Y.L. and Şen A. (2007). Spare parts inventory management with demand lead times and rationing, *IIE Transactions*, **39**, 879-898.

Köchel, P. and Nieländer, U. (2002). Defining optimal policies in multi-echelon inventory systems: The simulation optimization approach. In *12th Intern Working Seminar on Production Economics* **1**, 165-173.

Kranenburg, A.A. and van Houtum, G.J. (2004). A Multi-Item Spare Parts Inventory Model with Customer Differentiation. Department of Technology Management, Technische Universiteit Eindhoven. Working paper.

Lau, H.C., Song, H., See, C.T. and Cheng, S.Y. (2006). Evaluation of time-varying availability in multi-echelon spare parts systems with passivation. *European Journal of Operational Research* **170**, 91-105.

Lee, H.L. (1987). A multi-echelon inventory model for repairable items with emergency lateral transshipments. *Management Science* **33**, 1302–1316.

Lehmann, R. (2006). Customer benefits from LTP and O&M contracts. Retrieved: May 25, 2007, from URL: http://www.powergeneration.siemens.com/download/pool/1_Customer_benefits_from_LTP.pdf.

Li, H. and Xu, S.H. (2004). On the coordinated random group replacement policy in multivariate repairable systems. *Operations Research* **52**, 464-477.

Lu, H. and Yen, G.G. (2003). Rank-density-based multiobjective genetic algorithm and benchmark test function study. *IEEE Transactions on Evolutionary Computation*, **7(4)**, 325-343.

Marseguerra, M., Zio, E. and Podofillini, L. (2004). Multi-objective spare part allocation by means of genetic algorithms and Monte Carlo simulation. *Reliability engineering and System Safety* **1**(11). (in press).

Feldman, R. M., & Valdez-Flores, C. (1995). *Applied Probability and Stochastic Processes* (1st ed., p. 320). Pws Pub Co.

Medhi, J. (1982). *Stochastic Processes.* (p. 400). Wiley Eastern 1982.

Moinzadeh, K. and Lee, H.L. (1986). Batch size and stocking levels in multi-echelon repairable systems. *Management Science* **32**, 1567-1581.

Muckstadt, J.A. (1973). A model for a multi-item, multi-echelon, multi-indenture inventory system. *Management Science* **20**, 472-481.

Muckstadt, J.A. and Thomas, L.J. (1980). Are multi-echelon inventory methods worth implementing in systems with low-demand-rate items? *Management Science* **26**, 483-494.

Murthy, D.N.P. and Asgharizadeh, E. (1999). Optimal decision making in a maintenance service operation. *European Journal of Operational Research* **116**, 259-73.

Pal P., Das C.B., Panda A. and Bhunia A.K. (2005). An application of real-coded genetic algorithm (for mixed integer non-linear programming in an optimal two-warehouse inventory policy for deteriorating items with a linear trend in demand and a fixed planning horizon). *International Journal of Computer Mathematics* **82**, 163-175.

Power Engineering International, 2003. Parts pooling will save money for public power utilities. Retrieved: May 6, 2007, from URL: http://pepei.pennnet.com/display_article/173261/6/ARTCL/none/none/Parts-Pooling-Will-Save-Money-for-Public-Power-Utilities.

Rudi, N., Kapur, S. and Pyke, D. (2001). A two-location inventory model with transshipment and local decision making. *Management Science* **47**, 1668-1680.

Sherbrooke, C.C. (1968). METRIC: a multi-echelon technique for recoverable item control. *Operations Research* **16**, 122–141.

Sherbrooke, C.C. (1986). VARI-METRIC: improved approximations for multi-indenture, multi-echelon availability models. *Operations Research* **34**(2), 311-340.

Sherbrooke, C.C. (1992). Multi-echelon inventory systems with lateral supply. *Naval Research Logistics* **39**, 29-40.

Siemens AG Power Generation. (2004). Optimized spare parts logistics: Time is money. Retrieved: May 25, 2007, from URL: http://www.powergeneration.siemens.com/download/pool/PG_L2_2005_193.2en_Spares_Logisitcs.pdf.

Sleptchenko, A., van der Heijden M.C. and van Harten A. (2005). Using repair priorities to reduce stock investment in spare part networks. *European Journal of Operational Research* **163**, 733-750.

Sokhan-Sanj, S., Gaxiola, G., Mackulak, G. and Malmgren, F. (1999). A comparison of the exponential and the Hyperexponential distributions for modeling move requests in a semiconductor fab. In *Proceedings of the Winter Simulation Conference*, 774-778.

Srinivas N. and Deb K. (1994). Multiobjective optimization using nondominated sorting in genetic algorithms. *International Journal of Evolutionary Computation* **2(3)**, 221-248.

Triantaphyllou E., Kovalerchuk B., Mann L. and Knapp J. (1997). Determining the most important criteria in maintenance decision making. *Quality in Maintenance Engineering* **3**, 16-28.

Usher, J.S., Kamal, A.H. and Syed, W.H. (1998). Cost optimal preventive maintenance and replacement scheduling. *IIE Transactions* **12**, 1121-1128.

Vaughan, T.S. (2005). Failure replacement and preventive maintenance spare parts ordering policy. *European Journal of Operational Research* **161**,183-190.

Wang, Y., Cohen, M.A. and Zheng, Y.S. (2000). A two-echelon repairable inventory system with stocking-center-dependent depot replenishment lead times. *Management Science* **46**, 1441–1453.

Wong, H., Cattrysse, D. and van Oudheusden, D. (2005a). Inventory pooling of repairable spare parts with non-zero lateral transshipments time and relayed lateral transshipments. *European Journal of Operational Research* **165**, 207-218.

Wong, H., van Houtum, G.J., Cattrysse, D. and Van Oudheusden, D. (2005b). Simple efficient heuristics for multi-multi-location spare parts systems with transshipments and waiting time constraints. *Journal of the Operational Research Society* **56**, 1419-1430.

Wong, H., Cattrysse, D. and van Oudheusden, D. (2005c). Stocking decisions for repairable spare parts pooling in a multi-hub system. *Internationals Journal of production economics* **93-94**, 309-317.

Wong, H., van Houtum, G.J., Cattrysse, D. and Van Oudheusden, D. (2006). Multi-item spare parts systems with lateral transshipments and waiting time constraints. *European Journal of Operational Research* **171**, 1071-1093.

Yanagi, S. and Sasaki, M. (1992). An approximation method for the problem of a repairable-item inventory system with lateral supply. *IMA Journal of Mathematics Applied in Business and Industry* **3**, 305-314.

Zamperini, M. (2005). A Simulation Analysis of the Vari-METRIC Optimization Method for the U.S Coast Guard with a Part Failure Continuing Process Described by the Negative Binomial Distribution. Unpublished Master Thesis, Department of Industrial Engineering, Pennsylvania State University, University Park, Pennsylvania.

Zitzler, E., M. Laumanns and L. Thiele., 2001. SPEA2: Improving the strength Pareto evolutionary algorithm. Technical Report 103, Computer Engineering and Networks Laboratory (TIK). Swiss Federal Institute of Technology (ETH) Zurich, Gloriastrasse 35, CH-8092 Zurich, Switzerland.