DYNAMIC INVENTORY POLICIES FOR AEROSPACE SERVICE PARTS SUPPLY CHAIN

AGHIL REZAEI SOMARIN

School of Mechanical and Aerospace Engineering

A thesis submitted to the Nanyang Technological University in partial fulfilment of the requirement for the degree of Doctor of Philosophy

2014

Abstract

Service parts, which replace defective parts in aerospace repair facilities, are expensive. On the other hand, on-time fulfillment of demand for these parts is crucial as there are hefty financial penalties for aircraft schedule delays. Therefore, it is crucial to design an inventory and distribution system which minimizes the total cost of inventory investment. Recent changes in commercial aviation, e.g. substantial increase in fuel and labor costs and continuous growth of low fare carriers, have forced airlines to improve the efficiency of their MRO operations. As a result, it is not un-common for MRO operations to be outsourced to overseas service providers and for service providers to have more than one airline customer. All this has resulted in a complex and decentralized service parts logistics system.

Research so far has focused mainly on static decision making for service parts networks with a few warehouses. This research deals with generating dynamic inventory policies for larger service parts supply chains modeled as Markov Decision Processes (MDP). Capitalizing on the real-time information of parts in the resupply network, three types of decision are investigated: allocation of service parts to the bases (stock allocation policy), reallocation of service parts among the bases (stock reallocation policy) and emergency resupply of service parts from alternative source options (emergency resupply policy). Each policy is generated based on the optimal relative value function of the respective dynamic program modeled as MDP. It is shown that stock allocation policy can be characterized by a set of switching curves. Stock reallocation policy defines regions of imbalance. When the inventory levels enter these regions, reallocation of a service part is initiated. Emergency resupply policy can be defined by a set of limiting boundaries; when these boundaries are reached, it is optimal to fulfill the demand from alternative sources.

To overcome the common problem with MDP, "curse of dimensionality" in particular, heuristic techniques are proposed to approximate the optimal relative value function. The optimal relative value function for a single-base model is derived by solving the respective difference equation systems and is used as the basis for developing these techniques. The relative value function of a single-base model is proposed to be used for stock allocation policy generation. Aggregate queues are developed and utilized in approximating the optimal relative value function for stock reallocation. The optimal relative value function of an inventory system without safety stocks is derived and a difference operator is introduced. Using the proposed difference operator, the optimal relative value function for emergency resupply policy is approximated.

For each of the developed inventory policies, computational experiments are conducted to evaluate the optimality gaps and cost reductions. Computational experiments of each policy consist of a general problem set and series of sensitivity analysis. Based on the numerical results, proposed policies perform very close to the optimal policies. The models developed in this research could serve as a foundation to develop decision support systems to improve the efficiency and reduce costs in service parts supply chain management.

Acknowledgements

I would like to express the deepest appreciation to my principal supervisor, Prof. Chen Songlin. I have greatly benefited from his invaluable comments and feedbacks. Without his guidance and persistent help this thesis would not have been possible. I would also like to thank my co-supervisor, Dr. Lee Eng Wah. He has been greatly tolerant and supportive, throughout my PhD study. I am deeply grateful to my dear friends in MAE, Hassan, Sobhan and Ashkan. Discussions with them have always been illuminating, and their support and encouragement have been invaluable. I would like to show my greatest gratitude to my family for their persistent love and support. Finally, I would like to thank SIMTech and SINGA for their financial support.

Contents

List of Figures

List of Tables

Nomenclature

$\mathbf{A}(t)$	Number of part failures up to time step t
$\mathbf{B}(t)$	Number of finished repair jobs up to time step t
$\mathbf{c}$	Backorder cost per unit per time step
$C(\mathbf{x})$	Piecewise linear cost function
$\mathbf{e}'$	Demand fulfillment vector defined by the stock reallocation policy
$\mathbf{F}(t)$	Stock allocation decision vector at time step t
$\mathbf{E}(t)$	Emergency resupply decision vector at time step t
$\mathbf{G}^e(t)$	Emergency resupply decision vector at state $\mathbf{x}$
$\mathbf{G}^r(t)$	Stock reallocation decision vector at state $\mathbf{x}$
$h(\mathbf{x})$	Relative value function of the dynamic program
$\tilde{h}(\mathbf{x})$	Approximate relative value function
$J(\mathbf{x})$	Fluid value function
K	Number of bases

$\mathbf{R}(t)$	Stock reallocation decision vector at time step t
$\mathbf{S}$	Initial stocks allocated to the bases
T	Stock reallocation probability at each time step
$V(\mathbf{x})$	Value function of the dynamic program
$\mathbf{X}(t)$	Number of parts in the repair facility at time step t
$\mathsf{X}_{\mathbf{S}}$	State-space of the repair and allocation process characterized by the initial base-stock level $\mathbf{S}$
$\chi^o(\mathbf{x1}, \mathbf{x2})$	A trajectory generated by the optimal policy, starting from state $\mathbf{x1}$ and finishing at state $\mathbf{x2}$
η	Average cost
Λ	Demand probability at each time step
λ	Demand rate
μ	Repair rate
M	Repair probability at each time step
π	Stock reallocation cost
$\pi_e(\mathbf{x})$	Emergency resupply cost at state $\mathbf{x}$ under a particular emergency resupply policy e
π'	Emergency resupply cost

$\pi_r(\mathbf{x})$ Stock reallocation cost at state $\mathbf{x}$ under a particular stock reallocation policy r

$\Psi(\mathbf{x})$ Stock allocation policy

$\psi_{k_2}(\mathbf{x}_{k_1})$ switching curve of stock allocation between base k_1 and k_2 where $\mathbf{c}_{k_1} \leq \mathbf{c}_{k_2}$

$\tilde{\Psi}(\mathbf{x})$ Approximate stock allocation policy

$\Psi'(\mathbf{x})$ Stock reallocation policy

$\tilde{\Psi}'(\mathbf{x})$ Approximate stock reallocation policy

$\Psi''(\mathbf{x})$ Emergency resupply policy

$\tilde{\Psi}''(\mathbf{x})$ Approximate emergency resupply policy

τ Stock reallocation rate

Chapter 1

Introduction

1.1 Background

Capital goods like aircrafts, manufacturing facilities, power plants, may fail during their lifecycle. In these cases, the failed part needs to be identified and repaired/replaced within a reasonable time to assure minimum delay. To shorten the repair time, repair facilities prefer to hold reserve stocks of parts that are most likely to fail. These replacement parts are commonly called *"service parts"*. Clearly, service parts are crucial to the effective operations of any system. Without them, aircrafts will be grounded, manufacturing and construction processes will be delayed, and transportation and power supply systems will be disrupted.

Some of the failed parts can be economically repaired and be re-used, while others cannot be repaired or are not economical to repair. Thus service parts can be divided into *"Repairable"* and *"Non-repairable"* parts. Non-repairable parts are mostly cheap parts. Because of their low cost, batch ordering inventory policy is being used. Repairable items are usually expensive parts,[1] like those used in aerospace MRO, power plants, and manufacturing facilities. In this case, when a failure occurs, the failed part is replaced by a ready-to-use one, and the failed part is sent to be repaired at the repair facility. When the failed part cannot be repaired, a new part is purchased and placed in the repair facilities. Since service parts are expensive, it is crucial to design inventory control policies, which minimize the total value of inventory investment. On the other hand, critical industries like airlines, power plants and manufacturing facilities need to operate with minimum delay so the on-time fulfillment of demand for these parts is crucial. Hence, a cost effective service parts logistics system should be designed to minimize the total inventory and operations cost, while fulfilling customer's demand.

Depending on the distribution of the locations where demand occurs, service parts supply chain has different resupply network structures. In a single-echelon resupply network, all the inventory holding locations are supplied by a repair facility. The repair facility is responsible for repairing the failed parts and dispatching them to the inventory holding locations. In order to decrease the response time of the inventory holding locations to fulfill a demand, two-echelon resupply network is used. Two-echelon inventory structure were considered by early works in service parts supply chain for the military parts resupply systems [2]. In a two-echelon inventory system the upper level which is called a *"depot"* is responsible for repairing the complex and heavy parts and also resupplying ready-to-use parts to lower level warehouses (i.e. depot keeps stock and does the repairing

[1] e.g. repairable parts form 41% of the inventory investments in commercial aviation MRO [1].

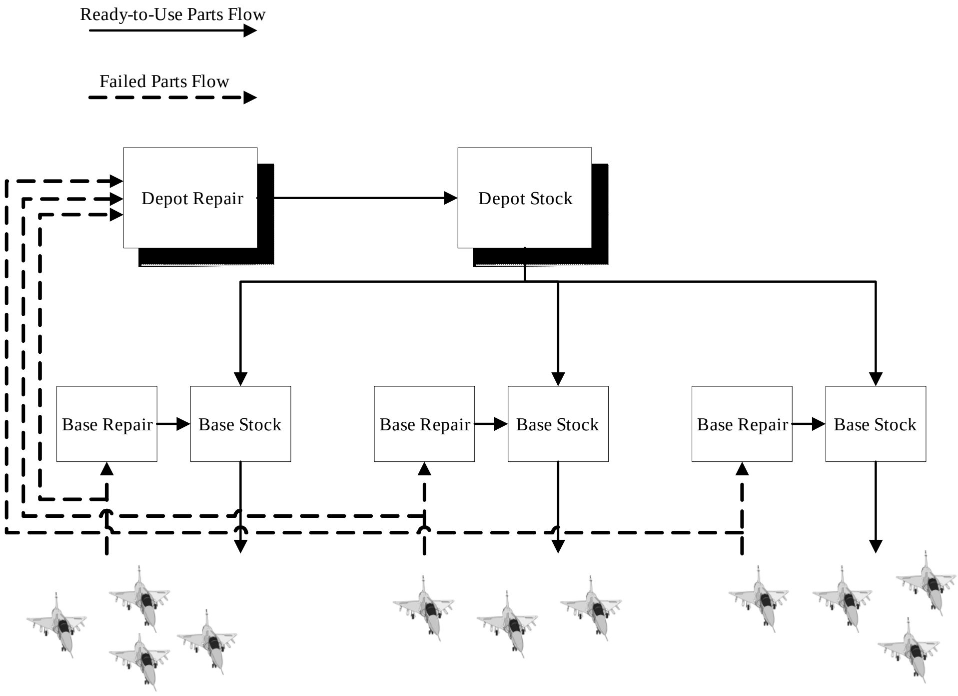


Figure 1.1: The depot-base resupply system

functions). Lower level warehouses or *"bases"*, provide ready-to-use replacement parts and it is also where simple parts are repaired. A regular military inventory system constitutes a depot and several bases. When an aircraft breaks down, technicians in the base identify the failed part and replace it with a ready-to-use part. If the part is repairable at the base, it is repaired and saved as a ready-to-use part. If the part cannot be repaired by the base, it is sent to the depot for repair. A ready-to-use part from the depot's reserve stock is sent to the base to maintain parts availability at base. See Figure 1.1 for parts flow in a depot-base resupply system.

With the advent of commercial service parts inventory systems, other network structures have emerged. When the customers are spread all over a geographical region, a multi-echelon structure is adopted to improve the response time. Figure

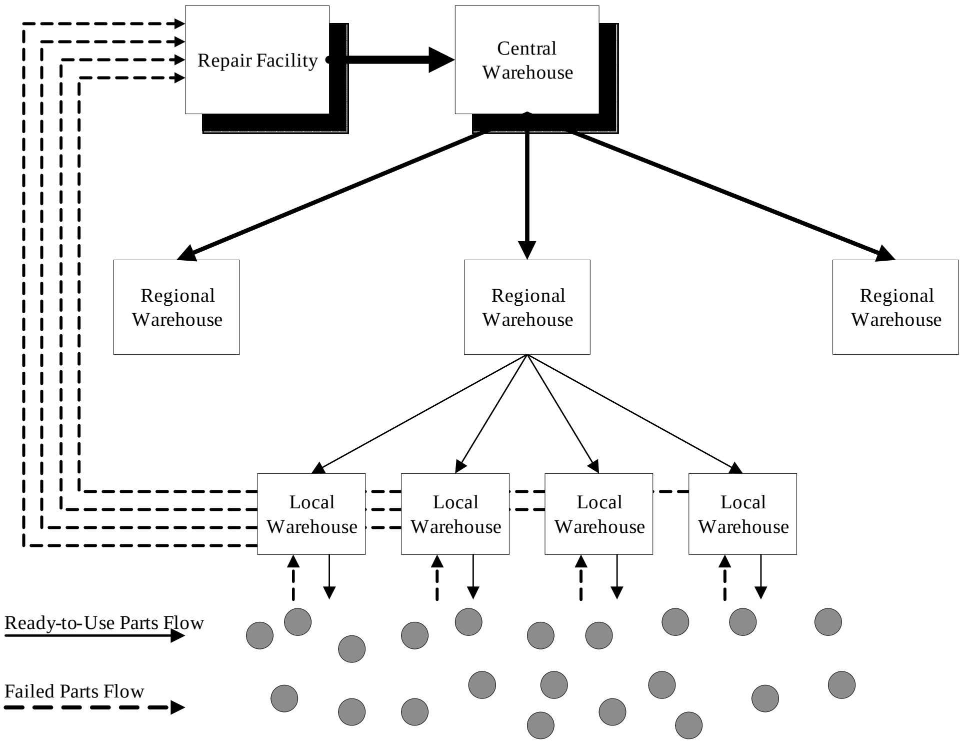


Figure 1.2: Three-echelon resupply system for repairable items

1.2 depicts a three-echelon parts resupply system. In most of the cases, the lowest level is just responsible for parts replacement while repair activities are done in the repair facility. This inventory structure is very common in service parts resupply systems with non-repairable items, in which the failed part is disposed upon replacement with a new part [3]. All three mentioned network structures are being used in service parts supply chain. However, based on recent developments in service parts inventory control, as the inventory policies become more flexible in fulfilling the demand, the need for multi-echelon network structure decreases.

In addition to the resupply network structure, inventory control policy plays a key role in the performance of the resupply system. Among the policies which can be utilized to improve the performance, one can mention ordering, stock allocation, reallocation and emergency resupply policies. One-for-one replenishment or $(S-1, S)$ is the popular ordering policy widely considered in the literature for slow moving items such as MRO service parts. In $(S-1, S)$ policy, whenever the inventory drops below the initial stock allocated to that base (S), the base asks a ready-to-use part from the higher echelon or the repair facility, in order to reach S again. For the proof of this policy's optimality for expensive parts, refer to [3]. Stock allocation policy involves the decision mechanism used by higher echelon or the repair facility to dispatch a ready-to-use part between lower echelon bases. In addition to the depot or repair facility, a base can be resupplied by neighboring bases, too. Bases are usually located closer together than with the depot; since the depot resupplies bases within the region it should be located close to them to shorten the transportation times between bases cf. the transportation times between a base and the depot. When a base is unable to fulfill a demand because there is no stock, or when the inventory level is too low, it is better to ask for a ready-to-use part from neighboring bases. Stock reallocation policy involves the lateral movement of parts between bases. Lateral movement can be physical or virtual. When the inventories of more than one customer have been pooled in a base, virtual movement of part refers to adjustment of the inventory share of each customer within the pool, while the total inventory level at base remains fixed. Stock reallocation policies developed here has been referred as *"lateral transshipment"* of parts between bases (physical movement of parts) and inventory rationing or multi-demand classes (in case of virtual movement of parts within the base) in the literature. Emergency resupply is another option considered in the literature. In the case of emergency resupply, a ready-to-use part can be available to a base in negligible time subject to an extra payment.

The payment can be regarded as a setup cost of an emergency repair server or the payment made to another MRO service provider, to exchange a failed part with a ready-to-use one.

1.2 Research Motivations

Competition in the airline industry has intensified over the years with steep increase in the cost of labor and fuel[1]. Additionally, low fare carriers continue to raise their market share with lower costs per Available Seat Mile [5]. In fact, the competition in the airline industry has been so fierce that major carriers have been forced to merge with their competitors to stay in business, e.g., Delta Air Lines-Northwest Airlines, Air France-KLM and Lufthansa-Swiss.

Such competitive pressure has also profound impact on airline operations, especially Maintenance, Repair and Overhaul (MRO). In fact, MRO is the largest component of an airline's operational cost (after fuel) [1]. The volume of the worldwide MRO market has grown from US$21B in 1995 to US$59.2B in 2013 and is projected to rise to US$85B by 2022 at 4.1% per annum [6]. As a result, MRO provides a major opportunity for reducing cost. For example, efforts in the last decade have resulted in reduction of the MRO supply chain inventory from US$3.6M to US$2.6M per active aircraft [5].

Airlines have responded to this trend in different ways. Low fare carriers, with their traditionally lean business model, have decided to outsource their maintenance to independent or third party maintenance providers. For example, many North American low fare airlines outsource their MRO to Central America. Ad-

[1]e.g. fuel price has increased by 500% from 2000 to 2008 [4].

ditionally, legacy airlines have been forced to either outsource their MRO, or to improve the efficiency of their internal MRO organization (e.g. by offshoring). Besides, those airlines that are still in the MRO business are looking to lower their maintenance costs by becoming a third party provider to other airlines. As an example, Lufthansa Technik has set up maintenance facilities for heavy checks in Hungary, China, Philippines and Malta. In addition, it has contracted with many airlines like Emirates, Air New Zealand, USA3000 Airlines and Aeromexico during recent years to provide MRO services for these airlines.

This trend (outsourcing, offshoring and servicing more than one customer) has resulted in a more decentralized and complex worldwide MRO supply chain structure, which is challenging to manage. In fact, MRO service providers typically serve many airline customers from all over the world each with a different expectation of the level of service. At the same time, in the interest of cost competitiveness, the MRO service providers need to minimize service parts stock within the network using more flexible inventory policies. From an extensive literature review, it is observed that most of the research in this area has so far focused mainly on decision making based on static information for networks with a few warehouses (as the models in the literature were initially developed for military applications). Hence, more dynamic and flexible inventory policies, which can be applied in a service parts logistics network as large as commercial aviation MRO resupply networks, is needed.

1.3 Research Objectives

The objective of this research is to develop dynamic service parts inventory policies for commercial aviation MRO. Specifically, three dynamic policies are considered in this research:

- *Stock allocation policy*: Based on the current inventory levels at the bases, where should the repaired part be sent?
- *Stock reallocation policy*: Based on the current inventory levels at the bases, is it necessary to transship a part from one base to another? If yes, from which base to which one?
- *Emergency resupply policy*: Based on the current inventory levels at the bases, if there is a source other than the repair facility or the regular higher echelon to fulfill the demand at the additional cost; when should one use that source; where should the part be sent?

1.4 Research Scope

This research is for repairable service parts logistics for the commercial aviation industry. Non-repairable parts is not covered in this study, nor will service parts for military or other applications. Compared with previous models, this research focuses on larger networks and also the capability of dynamic decision making. Throughout the thesis, it is assumed that developed policies are applied to a service parts inventory model with predetermined initial stock levels. In this study an inventory model for one part type is considered. Additionally, distribution network is assumed to be single echelon. That is, there is one repair facility with a repair server and several bases with safety stocks. However, in Chapter 7 two possible methods are presented which can be applied to extend the application of developed techniques to two-echelon structures.

1.5 Research Methodology

Research objectives are achieved by modeling the problem as Markov Decision Process (MDP). Afterwards, average cost infinite-horizon framework is considered for dynamic programming. Relative value function of the resulting dynamic program for single base model is obtained by solving the related difference equation systems. Optimal policies for a few number of bases is generated using the value iteration and policy iteration methods. For large number of bases, approximation techniques is proposed based on the optimal solution achieved for the single base model. Specifically, for stock allocation policy, an approximation method based on the relative value function of a single base model is proposed. Aggregate queues are constructed to address the problem of stock reallocation policy. Finally, capacitated aggregate queue is used to develop emergency resupply policy. Computational experiments are conducted to investigate the optimality gap between the proposed approximate policies and the optimal solution.

1.6 Thesis Outline

The remainder of this thesis is organized as follows: a review of the relevant literature is conducted in Chapter 2. In Chapter 3, the problem of stock allocation is addressed in the context of continuous-review system. Stock reallocation policy is presented in Chapter 4. Chapter 5 addresses the problem of emergency resupply. Chapter 6 concludes with a summary of the contributions and findings of this thesis, with some research area propositions for the future work.

Chapter 2

Literature Review

2.1 Introduction

Service Parts Supply Chain Management is the management of the flow of service parts, information and other resources, including repair facilities and human resources, to provide the right service part at the right location within a reasonable time to meet the customer requirements. Optimal service parts logistics should address issues like service parts stocking locations, inventory policies, repair facility location and capacity (i.e. number of repair servers) and the distribution and transportation networks.

Repairable service parts supply chain has been studied since the sixties. While earlier models were developed for military systems, recent work are commercially-focused, e.g. commercial aviation, power plants and consumer electronics. In this chapter the most relevant literature is reviewed. Existing literature on service parts supply chain can be classified into four categories based on the research problem: supply process, demand process, supplier/customer relationships and service parts.

2.1.1 Service Parts Supply Process

Supply process involves all the activities related to providing the ready-to-use parts to replace the failed parts of the faulty systems, including repairing the failed parts (for reparable items), distribution of the ready-to-use parts within the resupply network, control of the service parts inventories and cooperation with other suppliers to enhance the performance of the supply process by pooling inventories and thus, offsetting the backorder risk. Literature on service parts supply activities falls into four major subcategories: 1)distribution network, 2)control of service parts, 3)repair process, and 4)inter-supplier cooperation. Distribution network has been introduced briefly in Chapter 1. Service parts inventory control forms the major portion of service parts literature, which is discussed in detail in this Chapter.

Repair process involves modeling the repair and maintenance activities in the service parts supply chain. Joint inventory and maintenance modeling [7–16], repair prioritization [17–23], repair capacity determination [9, 24–27], and decisions regarding cannibalization [28–31] can be considered as major problems addressed in the literature. There can be several types of cooperative strategies in service parts supply chain depending on the strength of relationship and trust between part suppliers. In the most traditional strategy which is called solo strategy, each supplier serves just his own customers and there is no cooperation among suppliers. When there is a loose relationship between two suppliers, they may agree to loan spare parts to each other. If the relationship between suppliers is stronger and trust is high, assuming same rates of demand, they can pool the inventories and ask for lateral transshipment when one of them is out-of-stock. In this case, both suppliers can reduce their inventory levels and rely on the other party's inventory. This type of cooperation is called ad-hoc cooperation in the literature. Commercial pooling is another cooperation strategy addressed in the literature. In this strategy, two or more suppliers agree to cooperate based on a set of rules. These rules can include benefit sharing principles, response time to part requests from other parties, logistic connections between suppliers, inventory distribution between the bases, and the priorities in the stock-out situations. In this case, when a member uses one of the ready-to-use parts of the inventory pool, he is responsible to purchase a new part or repair the failed part and return it to the pool [32–38].

2.1.2 Service Parts Demand Process

Literature on service parts demand process includes investigations on the nature of demand and classification of service parts based on their demand occurrence frequency and quantities. Additionally, another stream of research effort has focused on a scenario where the demand is non-stationary because of the nature of the operating system. Hence, research efforts on service parts demand process can be classified into two subcategories: research on demand forecasting and research on variability of demand. The most common demand pattern considered in the literature for spare parts is "intermittent" demand. In this pattern, sequences of zero demand with occasional non-zero demands are observed [39–45]. While most of the literature assumes stationary demand process, this assumption is not always applicable to real world problems. A stream of research considers situations where demand process is non-stationary. These situations can be because of finite population size [46–48], population growth [49, 50] or population decay [51].

2.1.3 Supplier/Customer Relationship

While early models in service parts supply chain assume that repair service provider owns operative systems and inventories and is responsible for all the operations within the resupply network, recent publications consider the cases where the supplier for service parts is different from operative systems owners. In this case, contractual framework is the main focus of the literature [52–56]. Additionally, some publications consider the case where the supplier decides to stop production of out-dated parts and thus, asks inventory owners to place the final order for their service parts. End-of-life services is another area of research [57–59].

2.1.4 Service Part Structure

In addition to the above-mentioned categories, publications which focus on service parts and investigate their criticality for system operation [60–63], hierarchical structure [27, 64–66], and possibility of obsolescence of demand sources [67, 68] form the fourth category of the literature in service parts supply chain.

After providing a broad overview of the literature, the remainder of this chapter focuses on dynamic control of service parts. Most of the dynamic models developed for service parts inventory policies fall into two categories: stock allocation and stock reallocation policies. Another stream of research is involved with the emergency delivery of service parts. Little work has been done on this topic and most of the published papers have combined emergency delivery to the exiting models of stock allocation or reallocation. However, before investigating the dynamic models, a review of some of the basic models developed for early service parts inventory systems is presented, which are the basis for most of the dynamic inventory policies.

2.2 General Service Parts Inventory Models

Early models developed for service parts inventory management focus mainly on deriving performance measures for the resupply networks with particular initial stocking levels. Sherbrooke's [2] METRIC model is the first important work in the area of repairable items inventory models. METRIC is a two-echelon resupply network that was operated by US Air Force with a depot and a set of bases. The depot and bases hold inventory and conduct repair activities. METRIC's resupply system operates as follows:

When a part fails at a base, a ready-to-use part is withdrawn from the base's stock to replace the failed part. In the case of stock-out at base, backorder occurs. The part can be repaired either at a base or the depot. The decision on where to repair depends on the nature of the failure and availability of repair facilities at the base or depot. If the part is to be repaired at a base, the repaired part is saved as a ready-to-use part in the base. If the part cannot be repaired at the base, it is sent to the depot and a request for replenishment is made immediately. The depot sends one of its reserve stocks to the base and repairs the failed part and saves it as depot inventory. Also, if the depot has no stock, it backorders the demand from base. In the METRIC model, the following assumptions are made:

1. Part failure at bases follows the stationary Poisson process.
2. Repair times are statistically independent.
3. Each failure is caused by just one part.
4. The depot does not prioritize base requests, but deals with each on a FIFO basis.
5. Lateral transshipment is not allowed between bases.
6. All parts have the same criticality for aircrafts functionality.
7. All failed parts are assumed to be repairable (parts condemnation is not allowed).

Sherbrooke [2] mainly uses the expected backorder number to evaluate different inventory policies. Expected backorder number is the average number of backorders outstanding at a random point in time. Another performance measure to evaluate the stock policy is "Fill Rate". For a given stock level, the fill rate is the expected fraction of demand that can be fulfilled immediately by on-hand inventory.

Several authors have attempted to improve on METRIC. Simon [69] derives exact expressions for the stationary distributions of stock on hand, stock in repair, and backlogged demand at each facility of a two-echelon inventory model by analyzing the interval between a specific point in time and the time needed for the purchase of new parts and the repair of failed parts. In his model, the depot uses a one-for-one replenishment policy while bases use a continuous review (s, S) policy.

Graves [70] extends Simon's analysis to find both the mean and variance of backorders in depot and bases by approximating them to a negative binomial distribution. Using numerical results, Graves shows that his method gives more accurate approximations than METRIC proposed by Sherbrooke. However, while METRIC under-estimates the expected backorder, Graves' model over-estimates it.

Axsater [71] considers the interval between demand occurrence and its satisfaction as an Erlang distribution and proposes a method to find the optimal stocking policy directly by traditional cost functions. While his model is efficient when the cost is a nonlinear function of backorder or storage times (e.g. when the penalty cost is nonlinearly related to the backorder duration), it is inappropriate for cost functions that are expressed as nonlinear functions of the inventory levels or backorder size.

Taylor and Jackson [72] are the first researchers who modeled the service parts resupply system using queuing theory. They consider a three stage cyclic queue with stages: parts in operation, parts in repair and parts in stock. Using the birth and death process, they derive steady state probabilities for the parts in each stage. Using these probabilities, they obtain performance measures. One important limitation of their model is the assumption of a maximum of one backorder at a time.

Toft and Boothroyd [73] use the queueing model to predict the number of spare coal faces needed to continue production in the case of unforeseen closure of some coal faces. Assuming limited repair capacity, they derive the steady state probability of a given number of faces being out of action.

Mirasol [74] models the spare parts in service parts supply chain as customers in a cyclic queue. He assumes the commonly used hierarchical structure, and a one-for-one replenishment policy. Both interarrival times between failures and repairs are assumed to be exponentially distributed. Using balance equation, he finds the steady state probabilities for each stage within a cycle. He derives a new performance measure called "strategic unavailability" which is a function of both the probability and duration of backorders. He proposes a systematic procedure for determining the number of service parts as well as repair facilities. However, he assumes that the total number of parts within the system can't exceed the number of allocated service parts, which does not hold in the case of backordering, as in these situations the system would have more parts than the allocated level of service parts.

Gross et al. [75] model a single base service parts resupply system for marine gas turbine ships over multi-year planning horizon. They assume a finite component population, which changes during planning years as the demand rate. Also, with this method, they capture changing components reliability as older components are less reliable and thus have higher failure rates.

Gross et al. [76] model a large multi-echelon inventory system as a non-Jackson Markovian network with finite size. They approximate the steady-state behavior of the network by decomposing the network into smaller overlapping local subnetwork models and solving each subnetwork in isolation, iteratively updating the steady-state performance of each subnetwork until convergence. Gross et al. [77] present iterative solution methods to derive the steady-state probability distribution of a multi-echelon inventory system which is modeled as a Markov Process.

Albright and Soni [78] build two models. The first one is a model of single base resupply system using the continuous time Markov process with a multi-dimensional state space. Also, they use an exact algorithm based on aggregate/disaggregate methods, which require only the solution of a series of one-dimensional birth-death systems. The second model considers a two-echelon re-supply system with more than one base and base repair facility and one central repair depot. As the single base model, they use the aggregate/disaggregate method to develop a solution procedure. However, while for single base model they provide an exact solution, proposed solution for two-echelon structure is an approximation.

Wang et al. [79] consider a two-echelon inventory system with the same assumption as METRIC, only relaxing the assumption of i.i.d replenishment time. They show that for inventory systems with base dependent depot replenishment lead-time, such an assumption results in errors in estimating system performance.

Caglar et al. [80] consider a two-echelon, multi-item service parts inventory problem and develop a base-stock policy. They propose a heuristic algorithm to minimize the system-wide inventory cost subject to a waiting time constraint.

2.3 Stock Allocation Policies

Although many researchers have addressed the optimal inventory level and replenishment policies in service parts inventory system, only a few have focused on the problem of optimal stock allocation of the repaired parts to the bases.

Miller [81] addresses the dynamic stock allocation problem using the Markov decision process. He proposes a transportation look-ahead policy, in which the repaired part is sent to the base with the greatest marginal decrease in the expected backorder cost.

Federgruen and Zipkin [82] consider a dynamic ordering and allocation problem, where demand at bases has normal distribution. They approximate the myopic allocation policy and show that in this case, the problem can be simplified to a single location inventory problem.

Pyke [22] evaluates a simple dynamic stock allocation and repair prioritization policy using simulation and compares the performance with the first-come-first-serve policy. In his dynamic policy, the base with the highest number of aircrafts on the ground has priority over other bases in repair and stock allocation.

Bykkurt and Parlar [83] conduct a simulation study to compare two static and one dynamic policies. The policies that are evaluated are to send the part to the bases with fewer numbers of the parts in the system, to send the part to the base of origin and send it to the base with longest service parts shortage time. They conclude that the static policies perform similarly, while the dynamic policy outperforms both of them.

Caggiano et al. [84] formulate the problem of dynamic stock allocation with expedited shipment option and repair prioritization as a large scale linear program. They propose a greedy algorithm and rescale the stock numbers linearly in order to find the solution efficiently.

Barahona et al. [85] propose a method based on the weighted sum of the demand forecasts to allocate available stocks within a resupply network with flexible structure and non-stationary demand. Based on the simulation results they show that the flexible network, in which parts can be transhipped from every stocking point to another, outperforms the resupply networks with traditional hierarchical networks.

Tiemessen and van Houtum [86] investigate the possible decrease of aggregate downtime by smart repair job scheduling in a single-location multi-item model. They consider four heuristic scheduling rules, namely, modified equalization of run-out times, priority repairs, myopic allocation and shortest backorder times and show that the myopic policy outperforms other scheduling policies.

2.4 Stock Reallocation Policies

Stock reallocation involves movement of parts between bases in order to reduce the backorder risk. Researchers in this area have tried to improve the performance of the resupply network by introducing the capabilities such as emergency lateral transhipment and inventory rationing.

2.4.1 Emergency Lateral Transshipment

In multi-echelon inventory systems, transportation times between bases are shorter than the transportation times between the depot and bases. Thus in the case of stock-out, it is reasonable to fulfill a demand by an emergency lateral transshipment from neighboring bases. This approach seems to be reasonable considering the fact that usually repairable service parts are used in systems where uptime is crucial and lateral transshipment cost is significantly lower than backorder costs.

Numerous models have been developed for service parts resupply system with emergency lateral transshipment in the literature [25, 87–106] and most of them can be classified in different categories based on the configurations considered for inventory operations:

- **Centralized and Decentralized Decision Making**

  One of the main configurations which should be considered in the modeling of the spare parts resupply systems is the process of decision making. Some resupply systems are operated just by one repair service provider or the OEM of the product itself. In these cases, centralized decision making should be adopted to obtain the optimal inventory policy. In centralized decision making, all the information is available to the decision maker and he makes decisions that minimize the overall cost within the system. Decentralized decision making is appropriate for resupply systems with more than one repair service provider, who has the tendency to cooperate with each other. In this case, the decision maker only has access to information related to his own inventory operations. The main issue here is to decide when to borrow a part from cooperating service providers in order to maximize profit.

- **Unidirectional and Bidirectional Emergency Lateral Transshipment**

  Another important feature addressed by researchers is the direction of transshipments. While most of the literature adopted bidirectional lateral transshipments, some other assumed the possibility of lateral transshipment in just one direction [105]. These models are applied to resupply systems, in which, emergency lateral transshipment is possible in just one direction due to the nature of the transportation network or because the backorder cost is different at two locations, it is better to send a part from a location with lower backorder cost to the one with higher cost, but not vice versa.

- **Reactive and Proactive Emergency Lateral Transshipments**

  It has been assumed so far that a base asks for lateral transshipment when it is unable to fulfill a demand. Some of the researchers have adopted another scenario, in which the decision to send or receive a part by lateral transshipment is made proactively. In this case, decision for lateral transshipment request or supply is made either periodically at certain points in time or when events like demand occurrence triggers it. However in proactive lateral transshipment, the base in which demand has occurred may still have inventory on hand; the lateral transshipment just decreases the probability of stock-out, or in other words, the risk of going out of stock of all bases is balanced.

As an early work on emergency lateral transshipment Lee [87] considers a multi-echelon inventory system for repairable items that employs continuous monitoring of inventory similar to the METRIC model. Among the main assumptions of the model are identical bases, $(S-1, S)$ inventory control policy, first-come first-serve demand fulfillment and infinite repair servers. This model focuses mainly on cases where a fairly high service level is imposed at bases (above 70%). It divides the bases into groups within which emergency lateral transshipments are allowed. By using the results of [70] he finds the distribution of outstanding orders at the depot and consequently obtains the distribution of the number of orders outstanding at a particular pooling group by treating each group as a stocking location. Finally he derives the distribution of the number of orders outstanding at a base in the group. Besides, using a two phase method similar to that of [2], the optimal stocking levels at each location is obtained. By numerical results, it is shown that for the cases with higher levels of fill rate the approximation is accurate to 2 decimal places. However, for lower fill rate levels the model seems to be inaccurate. Besides, the identical bases assumption is unrealistic in practice.

Another seminal work in this area is by Axsater [88]. In this model, more emphasis is put on modeling the demand at a base correctly. Assumptions and decision rules in this model are similar to [87] to be able to compare the accuracy of approximations with it. An additional assumption is exponentially distributed repair times. The base for lateral transshipment supply is chosen randomly. It has shown by numerical results that this model is more accurate than Lee's model, when the proportion of lateral transshipment is relatively large. Besides, identical bases assumption has been relaxed in this model, which is more realistic. However, stock allocation decision making rule in this model is not optimal. While in the real world, when a repair task is finished, it is reasonable to send the part to a base with backorder rather than the original base where the part comes from, this model assumes the part to be sent to the original base. Consequently, there may be cases with stock on hand at a base and backorder at another base within the same group. Besides, the approach presented by Axsater seems to be appropriate for groups of two or three bases, since for higher number of bases within a group, modeling is difficult. This is a significant drawback when one wants to find the optimal number of bases to be established at each group using this approach.

Dada [89] incorporates priority shipment by specifying the system response when a base is out of stock, assuming a priority list is available for every base.

According to this model, when a demand occurs at a base, it is fulfilled by stock on hand. If the base is out of stock, the neighboring bases, depot and parts in transit from the depot to a base are looked for replenishment respectively. In this approach, an aggregate model of the system is first presented and then marginal probability distributions of bases within the system are approximated by disaggregation. It is assumed that all the parts at the lower echelon can be pooled. Because of the extensive geographical spread of multi-echelon systems, this is a restrictive assumption.

Sherbrooke [90] develops approximation model of a multi-echelon system with lateral transshipment between bases. Only depot repairs the failed parts. Based on his results, lateral supply is mostly efficient when the demand rates are low and transshipment times are short.

Kochel [91] models the behavior of inventory system by queueing theory. Considering the stock level as the number of servers at the queue station, inventory on hand as number of idle servers, replenishment time as service time and number of backorders as queue length, performance characteristics of a single location resupply system can be obtained. Also, using close queueing network, resupply systems with two location has been modeled.

Archibald et al. [92] consider a two-location inventory model with limited storage capacity for service parts. Lateral transshipment is allowed between locations. They model the problem as a Markov Decision Process, characterize the optimal policy as order up to policy and find it using the value iteration method.

Alfredsson and Verrijdt [93] propose a two echelon model consisting of a central warehouse supplying a number of local warehouses. When a demand occurs at a local warehouse, it is fulfilled by stock on hand . In the case of stock-out, demand would be fulfilled, either by emergency lateral transshipment, direct delivery from central warehouse or direct delivery from the plant. It was shown that for identical local warehouses, the model gives excellent results, while for non-identical local warehouses, it performs well. Also, sensitivity of performance measures regarding different distributions of replenishment times has been examined while they concluded that performance measures are almost identical for exponential, deterministic and lognormal replenishment times. Besides, by economical analysis it is shown that emergency lateral transshipment and direct delivery flexibilities can lead to significant cost reduction, while using only direct deliveries may cause dramatic increase in cost.

Grahovac and Chakravarty [94] present a decentralized model of a distribution center and several retailers, which is appropriate for commercial supply chains where retailers are independent and cost of purchase from distribution center is cheaper than neighboring retailers. Besides, this model allows for lateral transshipments when on hand inventory falls below a predetermined stock level (proactive lateral transshipment) and also considers the transshipment time in the model. Results show that introducing lateral transshipment may cause increase in inventory of retailers and decrease in distribution centers stock levels. Surprisingly, combining these two phenomena together may lead to higher level of inventory in the whole system when introducing lateral transshipment. However, it was shown that the system with lateral transshipment always outperforms the system without lateral transshipment.

Kukreja et al. [95] propose a model using Erlang loss system, in which lateral transshipment is requested from a location with lowest transshipment cost. Also, it is shown that there is no need for exponential replenishment times assumption, made by [88]. However, since lateral transshipment times have been ignored, this model seems to be appropriate only for the cases with negligible transshipment times between locations.

Jung et al. [25] model a resupply system with two levels of capacitated repair facilities and one level of stocking. In other words, the depot does not hold any inventory. Also, they provide an optimization method to find the optimal levels on service parts inventory in order to minimize the whole system's operating cost per unit time.

Wong et al. [96] propose a model based on machine-repair queueing model, where lateral transshipment times are considered. Also, upon repair completion, the part is sent to a base with backorder instead of the original base. Other assumptions include finite source of demand and ample number of repair servers. They use a two stage approach, where in the first stage all states with zero or positive inventory are examined. Among the states in stage 1, state zero is the state representing that no spare is available. This state is a compound state consisting of all the states examined in the second stage. Finally, performance measures i.e. fill rate, expected total downtime and expected number of lateral transshipments are obtained.

Wong et al. [97] consider a multi-item two-location inventory system with lateral and emergency shipments. A demand is fulfilled by emergency shipment when both of the bases are out of stocks. They propose a solution to minimize the cost subject to a waiting time constraint. The solution provides lower and upper-bound for the optimal total cost which is shown to be tight.

Zhao et al. [98] incorporate two types of decisions (decisions to supply and request a lateral transshipment) in their decentralized two-echelon model. In other words, in an extension to [94] where a critical inventory level is introduced for lateral transshipment request, another inventory level is considered. In this configuration, when a demand for lateral transshipment occurs, it is fulfilled only when the inventory level is higher than a specific point. This approach is mainly useful for decentralized decision making, where each base aims to minimize its own cost. In such models, special incentives should be set by cooperating or external units to enable lateral transshipment as fulfillment of these demands increases the probability of stock-out at the supplier base.

Kranenburg and van Houtum [99] consider a multi-item, multi-location, single echelon system, where locations are divided into main and regular warehouses. Since only a limited number of locations are needed to be equipped to supply lateral transshipment requests, this model is appropriate for the cases with high costs of installing transshipment equipments. Also, local warehouses are decoupled in two steps to provide approximation evaluations. First, regular warehouses are decoupled from mains, and in the second step, main warehouses are decoupled to individual warehouses to be analyzed separately. By numerical results, it is shown that most of the pooling benefits can be obtained by just equipping a small number of locations to provide lateral transshipments.

Kutanoglu [100] provides insights on the behavior of a two-echelon inventory system when inventory sharing is allowed between bases under the time-based service level framework. He finds that when inventory sharing is done only between two-bases, the full efficiency of the inventory sharing cannot be realized compared with group of three-bases. Additionally, when the demand in one of the bases increases, inventory sharing may lead to higher cost due to higher frequency of lateral transshipment.

Kutanoglu and Mahajan [101] model lateral transshipments with a two-echelon service parts resupply system, where infinite source of supply is assumed in higher echelon. Besides, it assumes bases are required to satisfy time-based service targets. In other words, system performance measures are defined in the context of specified time windows, e.g. 80% fill rate within four hours means that 80% of demands should be fulfilled within four hours either by stock on hand or lateral transshipment. The approximation approach used in this model, is similar to [88] and [93].

Van Utterbeeck et al. [102] examine three inventory configurations for both single and multi-echelon systems using simulation study: no flexibility, lateral transshipment and lateral transhipment with emergency deliveries. The model is appropriate for an environment in which the average waiting time for demand fulfillment is used as the performance measure. Their results show that increased resupply flexibility is nearly always beneficial and leads to significant cost savings. Besides, lateral transhipment technique is recognized as the main cost-reducing factor.

In another simulation study, Tiacci and Saetta [103] investigate the effect of two lateral transshipment approaches, namely, lateral transhipment based on availability and lateral transshipment for inventory equalization on mean supply delay. Their result show appreciable reduction of mean supply delay when lateral transshipments are allowed, while the latter approach slightly outperforms the first one.

Karsten et al. [104] investigate the stability of inventory pooling arrangement between independent decision makers, They address the issue of fairly distributing inventory costs over the participants when each party keeps a predetermined stocking levels as well as when the stocking levels are globally optimized.

2.4.2 Multi-Demand Classes

As mentioned before, inventory pooling between neighboring bases can dramatically reduce investment costs. In addition to lateral transshipment, inventory pooling can have another form, which also has been addressed in service parts supply chain literature. While in lateral transshipment with centralized decision making, a service provider manages his inventory located in more than one location; sometimes he manages the inventory of more than one customer at one or more locations. In this case, he can benefit from pooling the inventories allocated to each of his customers. When customers of a service provider have the same contract parameters (e.g. penalty rate, service level agreement), they can be treated as a single customer. However, in the real world, service provider offers his customers customized contracts. Thus, the service provider must manage the inventories of customers with different demand rates and contract parameters. To maximize his profit, the service provider can "share" or pool the inventories allocated for each customer. This is referred to as resupply systems with *"multi-demand classes"* and the decision making regarding the inventory management of different demand classes is called *"inventory rationing"*. Here, the issue is how to maximize the benefit of pooling the inventories of different customer classes by taking advantage of a customer's reserved stocks to fulfill a demand of another class and at the same time, meet the service level agreement of all the customers.

2.4.2.1 Rationing Decisions

Two types of decisions have been investigated by researchers in service parts resupply systems with multi-customer classes; these are:

1. **Demand Fulfillment**

   Demand fulfillment decision making in the management of inventories with multi-demand classes involves deciding if the demand from a particular class should be fulfilled or backordered. Since some customers have priority over others, it may be optimal to fulfill a demand from a higher customer class by the stocks allocated to lower class customers, or backorder a demand from a lower class customer in order to reserve stock for higher class customers. Besides, if the replenishment time is short, one may prefer to fulfill a demand from lower customer classes. Hence, time plays a crucial role in making demand fulfillment decisions.

2. **Stock Allocation**

   When a failed parts is repaired, the decision maker should decide on its allocation. One may think the optimal decision is to allocate the part to the higher class customer whose inventory has not reached its maximum level yet. But, decision making becomes difficult when there is a backorder from the lower customer class at the time of stock allocation. In this case priority could be given either to the stock reserved for a higher customer class or the backordered demand of the lower customer class, especially when a penalty rate is charged based on the waiting time. Another option is to allocate the part to the class which triggered the order.

2.4.2.2 Rationing Policies

In addition to the type of decisions to be made, the framework in which decision should be made is important. There are two common rationing policies:

1. Dynamic Rationing

   Dynamic rationing involves finding the optimal decision either at discrete fixed points of time or upon the occurrence of an event such as parts failure or repair completion. Approaches based on this decision making policy rely mostly on cost structure and use techniques such as dynamic programming. The main drawback of this approach is its inability to evaluate long run performance measures.

2. Rationing Using Critical Levels

   This approach determines threshold inventory levels to ease the process of optimal decision making during inventory operations. In contrast to dynamic rationing, this approach postulates long run performance measures can be derived. However, since it is mainly based on Markov processes and queuing theory, the modeling of many customer classes/locations is computationally intensive.

Nahmias and Demmy [107] consider a single location inventory model with two classes of customers. In their model, higher class represents those demand for parts without which the aircraft is not operative. They provide a rationing policy based on a critical level and compare the results with the inventory system without rationing. Ha [108] models a lot-for-lot production system with two customer classes as a queueing model. In this model, demands from higher class customer is fulfilled as far as there are stocks on hand, while demands from lower class customer may be backordered when inventory is low. A critical level is introduced for both production and inventory rationing, which is decreasing in the number of backorders of the lower class customer. When a demand from lower class occurs, it is fulfilled only when the inventory level is higher than this critical level. Also, when the inventory is lower than the critical level, system produces for higher class customer to reach the critical level. Assuming linear backordering costs, it has been proved that the above mentioned policy is optimal. However, modeling higher number of customers using this approach is difficult due to curse of dimensionality.

Dekker et al. [109] consider a spare parts inventory system where identical equipment has been installed in different machines, with different importance for the production process. Since in this case some of the demand has higher backorder cost, an inventory rationing policy is adopted. Introducing a critical level, all three options of stock allocation decisions explained before are investigated and approximations of service level for both critical and non-critical demands are derived.

de Vericourt et al. [110] discuss the stock allocation decision in a single-server, single product, make-to-stock queue with multiple demand classes, which is an extension to [108]. While [108] considers only two customer classes, this model is applicable for any number of customer classes.

Cattani and Souza [111] investigate inventory rationing policy of a production system with two customer classes. Higher customer class pays more and requires shorter lead times. They develop critical level rationing policy. When the inventory drops below the critical level, only demand from higher customer class is fulfilled.

Deshpande et al. [112] consider a static threshold-based rationing policy for an inventory model with pooling of inventory across two demand classes. The inventory system operates under a (Q, r) policy. They provide an efficient solution algorithm for computing stock controlling and rationing parameters and highlight the conditions where the proposed policy offers significant cost savings over traditional policies. A lower-bound on the cost of the optimal rationing policy is developed.

Fadiloglu and Bulut [113] propose a method for the exact analysis of inventory rationing problem. The method is based on the fact that if the continuous review inventory systems with backordering is sampled at multiples of supply lead-time, the state of the system evolves according to a Markov chain. Using the steady-state probabilities obtained by this Markov chain, underlying continuous review inventory system can be analyzed. Moreover, a recursive algorithm is developed to compute the steady-state probabilities.

As mentioned earlier, the main drawback of the models based on Markov processes is the curse of dimensionality, when the number of customer class increases. Vicil and Jackson [114] develop a method called Modified Bridge Algorithm-2 to reduce the complexity of the solution procedure caused by higher dimensions of Markov processes. Moreover, a heuristic algorithm is provided to find the steady state distributions. Finally, using an optimization scheme total system stock is minimized for a given set of service level constraints.

Arslan et al [115] consider a critical-level control policy for a single-product inventory system with multiple demand classes. The inventory system operates under a (Q, R) policy with rationing. They show the equivalence between this inventory system and a serial inventory system. A model for cost evaluation and optimization is developed and a computationally efficient heuristic is proposed.

Kocaga and Sen [116] present a single-echelon spare part distribution system with two demand classes. While classes considered by them have different priorities like the previous cases, demand fulfillment time is also different here; demand from higher class should be satisfied immediately whereas there is a fixed time allowed to fulfill a demand from the lower class. Modeling the system as a single-echelon inventory model, a static rationing policy based on critical level is proposed. Additionally, a lower bound approximation for service level is provided, which gives accurate approximations especially for higher levels of service. Moreover, they also examine the cases in which maximum benefit can be obtained by customer differentiation, namely, where lead-time or demand rate of lower class dominates the higher class and also when the difference between service level requirements of higher and lower class customers increase.

Kranenburg and van Houtum [117] discuss a single-location multi-item inventory system for spare parts with differentiated customer classes based on critical level policy. Demand which occurs when there is no inventory on hand is satisfied from an external source. They model the inventory system as a closed queueing network. By finding the steady-state probabilities for the number of parts in pipeline, service levels are derived. Optimal cost for the resulting optimization problem, which is a nonlinear integer programming problem is approximated by upper and lower bounds.

Teunter and Klein Haneveld [118] provide a critical level policy for two demand classes where the critical level is a function of remaining time until next replenishment. By characterization of a set of critical stocking times $L_1, L_2, \ldots$, corresponding critical inventory levels are obtained. For example, if the time until the next replenishment is shorter than L_1, no stock is reserved for higher class customer and if it is between L_1 and $L_1 + L_2$ only one stock is reserved for higher class customer and so on. It is shown that this policy outperforms the static critical level policies.

Enders et al. [119] propose an inventory model where customers are categorized into two groups of patient and impatient customers, based on whether they are willing to wait for demand to be fulfilled in case of stock out. They propose a critical level approach to differentiate two customer classes.

Alvarez et al. [120] propose selective emergency shipment of spare parts to differentiate customer classes. Two heuristic methods are presented based on local search and integer programming. It is shown that using the proposed policy combined with the critical level policy leads to significant cost reduction compared with the one-size-fits-all approach.

Tiemessen et al. [121] considers the problem of dynamic demand fulfillment in multi-location spare parts resupply network. In their model, part failure can be fulfilled from several locations by regular delivery or from an external source with an expensive emergency delivery. They propose a one-step look-ahead allocation rule and show that the proposed dynamic policy has small optimality gap.

Alvarez et al. [122] propose keeping dedicated stocks at customer sites as an approach to differentiate service levels of customers. In this case, the single location model changes to a two-echelon model with a central warehouse and several local warehouses at customer sites. They consider the system under backordering and the use of emergency shipments. Using numerical experiments, they show that, same as critical level policy, service differentiation using dedicated stocks outperforms the static policy of providing all customers with uniform service.

2.5 Emergency Resupply Policies

In addition to the emergency delivery models reviewed in stock allocation and reallocation sections ([93, 97]), a few models have been developed focusing on the emergency delivery of spare parts.

Moinzadeh and Schmidt [123] develop an approximate model for a single location inventory model with $(S-1, S)$ ordering policy. Emergency orders are allowed in their model with an additional cost. The model takes advantage of the information about the age of outstanding orders and considers both backorder and lost sales case.

van der Heijden [124] develop a multi-echelon multi-indenture spare parts model in which the repair and transportation time can be reduced at additional costs. They propose an optimization heuristic for the cost trade-off between throughput time reduction and spare parts inventory. Numerical experiments indicate a 20 % cost reduction on average.

Ozkan et al. [125] provide accurate approximations of the performance in a two-echelon inventory model with extra option of demand fulfillment via emergency delivery from the central warehouse (in case of stockout at local warehouse) or delivery from the repair facility (in case of stockout at the local and central warehouses).

Axsater et al. [126] consider a two echelon inventory system where the lower echelon warehouses are supplied by a support warehouse in case of stock out. They present a model for cost evaluation and optimization of the reorder points under service level constraints.

2.6 Summary and Research Gaps

The seminal work of Sherbrooke [2] is the pioneer research effort in this field. In his paper, Sherbrooke proposed a method to approximate performance measures for a specific inventory policy used by the US Air Force. Subsequent works [69, 127, 70, 64, 128] concentrated on improving the accuracy of these performance measures. All these models assumed that a warehouse is supplied only by upper-level warehouses or by the repair facility. Lee [87] however considers a case in which the warehouse could borrow parts from a neighboring warehouse when there is a stock-out. While [87] was an extension of Sherbrooke's model, subsequent works modeled the service parts re-supply network as Markov processes. Because of the flexibility of Markov processes in modeling different demand rates during inventory operations, this approach was able to relax some of the assumptions made by earlier works and thus provided more accurate performance measures. The same approach was the basis for subsequent research efforts, in which demand from several customer classes could be fulfilled optimally by reserving stock for higher customer classes. Arising from the literature review, the author identified the following two gaps are identified to be the most pressing in light of present-day aviation MRO:

- Researchers in the past analyzed aviation parts re-supply networks of limited size. As aviation operations today are characterized by a large number of bases involving lateral transshipment, the methods they adopted are inadequate today. Although most of the researchers have applied methods based on Markov processes, they have not been able to find closed-form solution for a large number of inventory holding locations with lateral transshipment. They have modeled each base separately, while some of them assumed that information on the demand rate from neighboring locations is available [94, 98] and yet others have adopted an iterative solution process to find the mentioned lateral demand rate [88, 93, 101]. However, the first assumption is simplistic and is not appropriate for systems with centralized decision making. As for the second approach, the modeling process is rather complicated for a large number of bases; in fact, for every change in the number of bases, a new model (based on the same approach) is needed.

- Most of the past research assumed that timely information about the stock levels in the bases and depot is not available. However, recent parts tracking technologies such as RFID have made available real-time information for decision-making. There is a lack of literature that explicitly incorporates such real-time information.

After identifying these two main research gaps from the literature, dynamic service parts inventory policies for commercial aviation MRO are developed addressing these research gaps. Specifically, three dynamic policies are considered in this research; stock allocation, stock reallocation and emergency resupply policies. Proposed dynamic policies utilize the real-time information regarding the inventory levels at each base for decision making and can be implemented in the inventory systems with large number of bases.

Chapter 3

Stock Allocation of Service Parts to the Bases

3.1 Introduction

In this chapter, base-stock provisioning and repaired part allocation problem for a single part type, multi-location service parts inventory model is considered. The problem is modeled as a Markov Decision Process (MDP) and the dynamic program of the problem is provided. Some of the characteristics of the optimal stock allocation policy are investigated and it is shown that finding the optimal policy for one of the two stock allocation decisions (base-stock provisioning and repaired part allocation to the bases) leads to the optimal decision for the other. Additionally, a heuristic method for the stock allocation is proposed based on a modification of the relative value function of a single base where the average cost impacts the policy more as the number of bases in the inventory system increases. The performance of the proposed method is evaluated by numerical experiments, where it is compared with the optimal and myopic policies.

The remainder of this chapter is organized as follows: The problem under study is described and model formulation is presented in Section 3.2. In Section 3.3.1, the problem of optimal stock allocation is addressed and the heuristic method is presented. Numerical results of the experiments are presented and discussed in Section 3.4. Section 3.5 concludes with a summary of the contributions and findings of this chapter.

3.2 Problem Description and Model Formulation

In this section the inventory system configurations is explained, the notations are introduced, the assumptions are discussed and the stock allocation is formulated using MDP.

3.2.1 System Description

Consider a repairable service parts inventory system with a central repair facility called depot and K service parts holding locations called bases. Parts fail with Poisson distribution and base k faces demand for the ready-to-use parts with the rate of λ_k. After a part failure, failed part is identified and replaced with a ready-to-use part from base inventory. If no ready-to-use part is available at the base, the demand is backordered and fulfilled as soon as a ready-to-use part becomes available. The failed part is sent to the repair facility. Central repair facility repairs the part with exponential repair time at the rate μ and sends it to

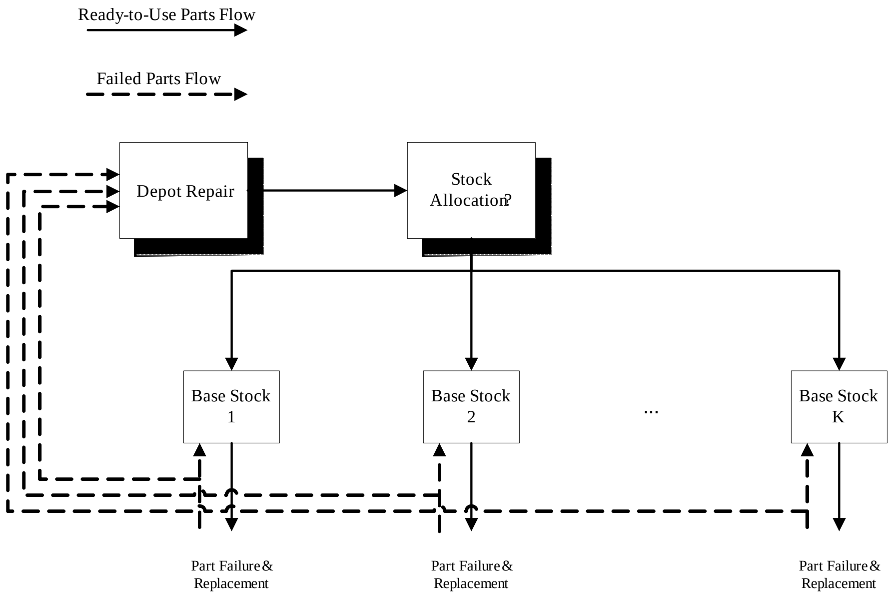


Figure 3.1: Service parts inventory system and operations

one of the bases. It is assumed that μ is identical and mutually independent for all of the bases and the transportation times from the repair facility to the bases are negligible. An inventory system with one part type and several customer classes is considered, namely, backorder costs at bases are non-identical. Figure 3.1 illustrates a graphical representation of the inventory network and service parts movement within the resupply network.

At the beginning of operation, initial stocks are provisioned and allocated to each base. During the operation, when a part is repaired, repair facility allocates the ready-to-use part to the bases. The objective is to find a stock allocation policy which minimizes the expected backorder cost subject to an initial investment budget in the form of the total number of service parts to be allocated initially to the bases. To achieve this objective, one should address a tactical level decision (initial stock provisioning) and an operational level decision (stock allocation policy after repair).

3.2.2 Notations and Assumptions

It is assumed that the parts at the bases fail with Poisson distribution. Poisson demand distribution is inline with most of the literature. The reason for this assumption is that part failure probability is commonly assumed to have binomial distribution and since the number of parts operating in a system is assumed to be large, the demand process at each base can be approximated by Poisson distribution. Earlier models of service parts inventory system simply assumed infinite number of repair servers at the repair facility. Using this assumption and based on Palm's theorem, they were able to show that the steady-state number of parts at the repair facility depends only on the average repair time and is independent of the repair time distribution. This leads to extensive use of exponential repair times in the literature to be able to model the inventory system as a Markov process. While it is assumed that there is a single repair server in the repair facility, exponential repair time assumption is made to model the stock allocation decision as a Markov Decision Process (MDP). Additionally, it is assumed that server utilization at the repair facility is less than one. Negligible transportation time seem to be reasonable, considering the fact that in the real world, repair time forms the large part of the whole replenishment time and transportation time from depot to base is proportionally much shorter than repair time.

In order to avoid the computational complexity of the stock allocation solution, single repair server is assumed to operate in the repair facility. Initial base-stock provisioning is constrained by an initial investment budget, i.e. the total number of purchased service parts is limited and it is assumed that during the inventory operation, it remains constant. It is shown that the solution to the optimal initial base-stock provisioning in this case is similar to the solution of optimal stock allocation of repaired parts to the bases. Moreover, it is assumed that there is no cost incurred while the bases have inventory on-hand and a penalty cost $\mathbf{c}_k$ is charged per unit per time when the base k has backorder.

Throughout the chapter, x denote a variable, while $\mathbf{x}$ denote a vector of appropriate size. $\mathbf{x}_k$ denote the kth element of the vector $\mathbf{x}$ and $\bar{\mathbf{x}}$ denote the sum of the elements of vector $\mathbf{x}$. $\mathbf{e}_k$ denote a vector of appropriate size with a one at kth element and zero at other elements. Additionally, for two vector of the same size $\mathbf{x}$ and $\mathbf{y}$, $\mathbf{x} \prec \mathbf{y}$ denote the element-wise comparison, that is if $\mathbf{x} \prec \mathbf{y}$, then for each k, $\mathbf{x}_k < \mathbf{y}_k$.

3.2.3 MDP Formulation

The inventory system described in Section 3.2.1 can be modeled as a queueing model of parallel queues with a server, where queues and jobs inside them represent the bases and failed parts sent to be repaired, respectively. In this manner, stock allocation decision can be modeled as the decision on which queue should be served. To formulate this inventory system, infinite time horizon model transformed into a discrete-time Markov Decision Process is considered and the modeling approach presented in Meyn [135] is used. Let $\mathbf{S}_k$ denote the number of service parts initially allocated to base k. Let $\mathbf{X}_k(t)$ denote the number of parts from base k in the repair facility at time t. Let $\mathbf{F}(t)$ denote stock allocation decision vector, where $\mathbf{F}_k(t) = 1$ means if a part is repaired at time t, it is sent to the base k. Clearly, $\mathbf{F}_k(t) \in \{0, 1\}$, $\sum_{k=1}^{K} \mathbf{F}_k(t) \leq 1$ and $\mathbf{F}_k(t) = 0$ whenever

$\mathbf{X}_k(t) = 0$. $\mathbf{X}_k(t)$ can be considered as queue length process such that at each time step, the number of parts in repair from base k increase by 1 with probability Λ_k and decrease by 1 with probability $\mathbf{F}_k(t-1)M$, where probabilities Λ_k and M can be defined as follows:

$$\Lambda_k = \frac{\lambda_k}{\mu + \sum_{m=1}^{K} \lambda_m} \quad (3.1)$$

$$M = \frac{\mu}{\mu + \sum_{m=1}^{K} \lambda_m} \quad (3.2)$$

The number of failed parts at the repair facility can be treated as a queue length process defined by the following recursion:

$$\mathbf{X}(t) = \mathbf{X}(t-1) - B(t)\mathbf{F}(t-1) + \mathbf{A}(t) \quad (3.3)$$

where the i.i.d process $(\mathbf{A}_k(1 \leq k \leq K), B)$ has marginal distribution defined by

$$\mathsf{P}\{(\mathbf{A}_1(t), \mathbf{A}_2(t), ..., \mathbf{A}_K(t), B(t)) = \mathbf{e}_i\} = \mathbf{p}_i, \quad (3.4)$$

with $\mathbf{p}_i = \Lambda_i$ for $i = 1, 2, ..., K$ and $\mathbf{p}_{K+1} = M$. In this case, $(\mathbf{X}, \mathbf{F})$ is a Markov Decision Process, where the stock allocation decision $\mathbf{F}(t)$ is defined based on the number of parts in repair from each base. Assuming the current state is $\mathbf{x} = \mathbf{X}(t)$, the stage cost at time t is:

$$C(\mathbf{x}) = \sum_{k=1}^{K} \mathbf{c}_k[\mathbf{x}_k - \mathbf{S}_k]^+. \quad (3.5)$$

3.3 Stock Allocation Policy

Let $h^*(\mathbf{x})$ and $\eta^*(\mathbf{S})$ denote the relative value function for the dynamic program and the steady-state average cost of the system under the optimal stock allocation policy, respectively. The average cost is independent of the initial state $\mathbf{x}$ and the dynamic programming optimality equation [135] can be defined as follows:

$$\eta^*(\mathbf{S}) + h^*(\mathbf{x}) = \min_{f \in \mathsf{F}(\mathbf{x})} [C(\mathbf{x}) + (\sum_{k=1}^{K} \Lambda_k h^*(\mathbf{x} + \mathbf{e}_k) + M h^*(\mathbf{x} - \mathbf{e}_f)]. \tag{3.6}$$

Without loss of generality it can be assumed that the state of origin is 0, or there is no failed parts being repaired or waiting to be repaired at the central repair facility. Let $\Phi(x)$ denote the stock allocation policy when current state is $\mathbf{x}$. In this case, optimal stock allocation policy can be defined based on the optimal relative value function as follows:

$$\Phi^*(\mathbf{x}) = \arg \min_{f \in \mathsf{F}(x)} (h^*(\mathbf{x} - \mathbf{e}_f)). \tag{3.7}$$

Optimality equation can be solved either by using the difference equation systems or iterative optimization techniques available for dynamic programming. Using the difference equation systems, analytical solutions can be derived for the optimal relative value function of problems with very simple structure. Optimization techniques consist of recursive algorithms such as value iteration or policy iteration methods. In these methods a sequence of functions are defined and updated recursively. For example, in the value iteration algorithm, the sequence of value

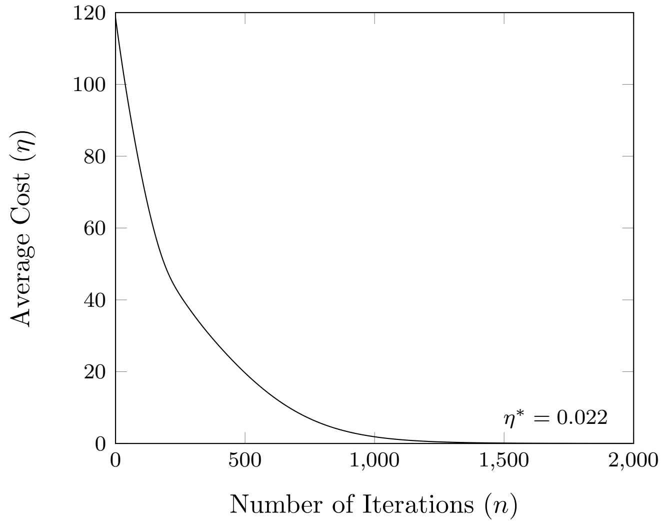


Figure 3.2: Average cost convergence in the value iteration method

functions ($V_{i+1}(\mathbf{x})$) are defined as follows:

$$V_{i+1}(\mathbf{x}) = \min_{f \in \mathsf{F}(\mathbf{x})} [C(\mathbf{x}) + (\sum_{k=1}^{K} \Lambda_k V_i(\mathbf{x} + \mathbf{e}_k) + M V_i(\mathbf{x} - \mathbf{e}_f)] \tag{3.8}$$

After running the algorithm for many steps, $V_{i+1}(\mathbf{x}) - V_i(\mathbf{x})$ converges to the optimal average cost. Subsequently, the optimal control policy is such that minimizes:

$$\Phi^*(\mathbf{x}) = \arg \min_{f \in \mathsf{F}(\mathbf{x})} (V_i(\mathbf{x} - \mathbf{e}_f)). \tag{3.9}$$

Stock allocation policy depicted in Figure 3.3 has been obtained using the value iteration method. Demand rates are $\lambda_1 = 2$ and $\lambda_2 = 1$, the repair rate is assumed to be $\mu = 3.75$ and backorder costs are $c_1 = 1$ and $c_2 = 2$. Initial stocks allocated to the bases are $S_1 = 16$ and $S_2 = 8$. State space is divided into two regions: send the repaired part to the base 1 and send the repaired part to the base 2. As can be seen, two regions have been separated using a switching curve.

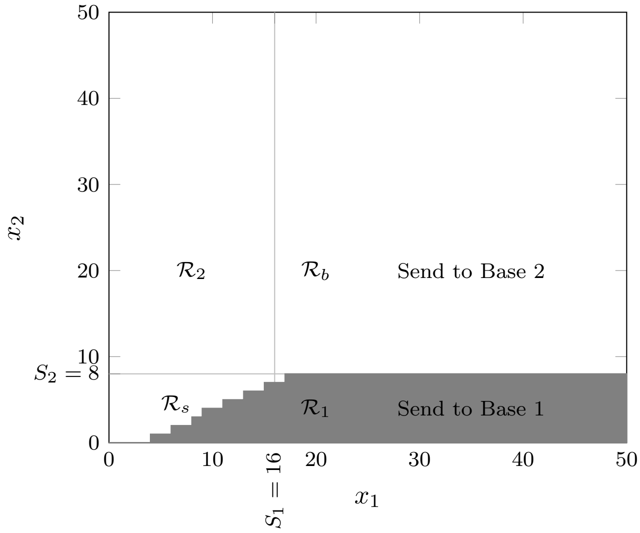


Figure 3.3: Stock allocation policy for a two-base model

Region $\mathcal{R}_b$ represents the case of both bases having backorders and based on the $c\mu$ rule, it is optimal to send the part to the base with higher backorder cost (See [136, 137]). $\mathcal{R}_s$ represent the case of both bases having inventory on-hand. In this case, a switching curve dictates the stock allocation policy. Regions $\mathcal{R}_1$ and $\mathcal{R}_2$ represent the cases of having backorders at base 1 and 2, respectively, while the other base has inventory on-hand and it is optimal to send the part to the base with backorder. However, this is not always the case; in some other cases, it might be optimal to send the part to the base with higher backorder cost while it has inventory on-hand, despite existing backorders in the base with lower backorder cost. In fact, there is a critical level for the base with higher backorder cost. Whenever, its inventory level drops below that critical level, it is optimal to send the repaired part to that base, despite the existing backorders in the base with lower backorder cost. Figure 3.4 illustrates the stock allocation policy for an example with $\lambda = [1, 2]$, $\rho = 0.8$, $c = [1, 2]$ and $S = [8, 16]$.

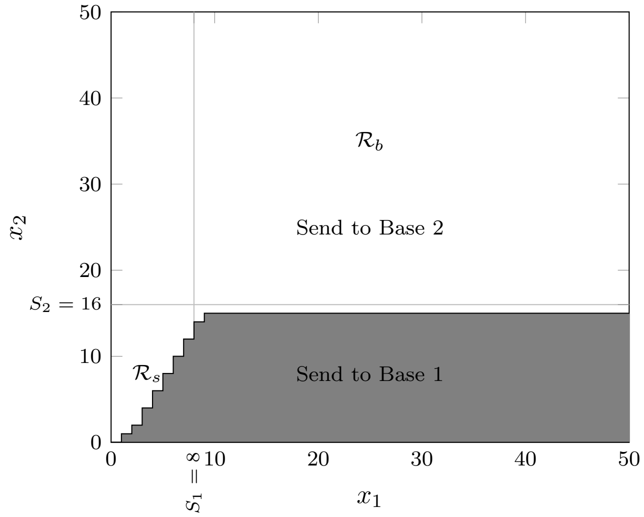


Figure 3.4: Stock allocation policy for a two-base example with critical level

As can be seen in Figure 3.4 when there is backorder at base 1, switching curve constitutes a straight line. De Vericourt et al. [138] characterize the structure of the optimal policy in this region using sample-path comparison and dynamic program. Let $\psi_{k_2}(\mathbf{x}_{k_1})$ denote the switching curve of stock allocation between base k_1 and k_2 where $\mathbf{c}_{k_1} \leq \mathbf{c}_{k_2}$. Based on their result, when $\mathbf{x}_{k_1} > \mathbf{S}_{k_1}$, switching curve can be found as follows:

$$\psi_{k_2}(\mathbf{x}_{k_1}) = \mathbf{S}_{k_2} - \lfloor \frac{\ln(\mathbf{c}_{k_1}/\mathbf{c}_{k_2})}{\ln(\lambda_{k_2}/\mu)} \rfloor. \tag{3.10}$$

In the value iteration method, a control policy is obtained after running the algorithm for many steps. For infinite capacity queues such as the queue length process considered here, using the value iteration method is impractical unless the capacity of each queue is limited and the approximate value functions are found. Besides, the algorithm needs to run for many steps in order to converge to the optimal average cost (Figure 3.2). Additionally, computational time and the amount of memory needed to keep the value function increases exponentially with the increase in the number of bases. To address the limitations of these approaches, researchers have suggested heuristic methods. Tiemessen and van Houtum [86] consider four such heuristic rules and compare their performance under different scenarios. Generally, heuristic rules are either static or based on some functions defined only on one dimension, i.e. for each base these functions can be calculated separately. The allocation rule chooses the base with lowest or highest function value depending on the rule.

3.3.1 Some Properties of the Optimal Stock Allocation Policy

In this section some properties of the optimal stock allocation policy is presented and it is shown that finding the optimal initial base-stock levels for a series of initial investment budgets starting from zero to $\mathbf{S}$, leads to obtaining the optimal stock allocation switching curve. Afterwards, a heuristic method is proposed to approximate the initial stock levels and the stock allocation switching curve.

Stock allocation switching curves can be defined more precisely. Within the state-space, state $\mathbf{x}$ is on the stock allocation switching curve, if and only if, for each $k \in K$, $\mathbf{F}_k(\mathbf{x} + \mathbf{e}_k) = 1$. If the current state $(\mathbf{x})$ is on the stock allocation switching curve and there is a failure at base k and a failed part is repaired before any further failures at other bases, it is sent to base k.

Let $\mathbf{y}$ denote any arbitrary state, where $\mathbf{y} \succeq \mathbf{x}$ and $\mathbf{y} \neq \mathbf{x}$. Let $\chi^o(\mathbf{x1}, \mathbf{x2})$ denote a trajectory generated by the optimal policy, starting from state $\mathbf{x1}$ and finishing at state $\mathbf{x2}$. Based on the structure of the optimal stock allocation switching curve, the trajectory $\chi^o(\mathbf{y}, \mathbf{0})$ passes state $\mathbf{x}$ at least once. For more discussion on this subject refer to De Vericourt et al. [138].

Let $\mathsf{X}_{\mathbf{S}}$ denote the state-space for repair and allocation process characterized by the optimal initial base-stock level $\mathbf{S}$. Let $\mathbf{S}'$ denote any optimal initial base-stock level with lower investment budget, that is $0 \succeq \mathbf{S}' \succeq \mathbf{S}$. State $\mathbf{z}$ is defined as $\mathbf{z} = \mathbf{S} - \mathbf{S}'$ and $\mathsf{X}_{\mathbf{S}'}$ is called a subspace of $\mathsf{X}_{\mathbf{S}}$. By definition, for any arbitrary state $\mathbf{y}$ in subspace $\mathsf{X}_{\mathbf{S}'}$, trajectory $\chi^0(\mathbf{y}, \mathbf{z})$ is minimum. It is shown that if state $\mathbf{x}$ is on the stock allocation path and $\sum_{k=1}^{K} \mathbf{x}_k = \sum_{k=1}^{K} \mathbf{z}_k$, then $\mathbf{x} = \mathbf{z}$. Optimal backorder cost for the trajectory $\chi^o(\mathbf{y}, 0)$ can be represented as the sum of the backorder costs for trajectories $\chi^o(\mathbf{y}, \mathbf{x})$ and $\chi^o(\mathbf{x}, \mathbf{0})$. Because of the Markovian property, two costs are independent and both trajectories $\chi^o(\mathbf{y}, \mathbf{x})$ and $\chi^o(\mathbf{x}, \mathbf{0})$ are minimum. Hence, $\mathbf{x} = \mathbf{z}$ and if a state is on the stock allocation switching curve in the subspace $\mathsf{X}_{S'}$, its respective state in state-space X_S is on the stock allocation switching curve, too and vice versa.

It follows from this property that the optimal initial base-stock level is on the stock allocation switching curve. Furthermore, in order to find the optimal initial base-stock level, one can start from $\mathbf{S}' = 0$ and find the k which minimizes the average cost for the inventory system with $\mathbf{S}' = \mathbf{S}' + \mathbf{e}_k$ base-stock level, until all the available service parts are allocated. This approach gives the optimal initial base-stock level and the optimal stock allocation switching curve, at the same time. However, deriving the average cost for the inventory systems with large number of bases is intractable or computationally intensive. A heuristic method based on simplistic assumptions is proposed, which approximates the optimal policy accurately.

3.3.2 Approximate Stock Allocation Policy

Consider an inventory system with two bases, $\mathbf{S}$ initial base-stock levels and $\mathbf{S}'$ on-hand inventory, while $\mathbf{x} = \mathbf{S} - \mathbf{S}'$ is on the stock allocation switching curve. Based on the definition of the switching curve, if the next event is a part failure in one of the two bases, repair facility allocates its whole capacity to repair a part and send it to the base with part failure at state $\mathbf{x}$. This policy is optimal if until the next time the state $\mathbf{x}$ is revisited, no part failure takes place at the other base. Based on this assumption, two bases can be considered and analyzed separately. Myopic policy introduced by [139] takes into account the rate of increase in the expected instantaneous cost by looking ahead over the repair time. It then allocates the repaired part to the base with smaller rate of increase. Hence, this assumption is inherent in myopic policy. It is proposed that when the current state is $\mathbf{x}$ or on-hand inventory is $\mathbf{S}'$, the next repaired part is sent to the base with highest increase in the relative value function resulting from a part failure at that base. In other words, while in the myopic policy the set of events and their respective cost until the next repair job is done is considered, in the value function based method, the set of events until the next revisit of the current state is considered. To derive the approximate stock allocation switching curve, the relative value function for two bases with $\mathbf{S}'$ base-stock level at state 1 (i.e. $h_k(1)$) needs to be derived. Based on the average cost optimality equation:

$$\eta_k(\mathbf{S}'_k) + \lambda_k h_k(0) = \lambda_k h_k(1). \tag{3.11}$$

By definition $h_k(0) = 0$ and $h_k(1)$ can be derived easily as follows:

$$h_k(1) = \frac{\eta_k(\mathbf{S}'_k)}{\lambda_k}. \tag{3.12}$$

where the average backorder cost is as follows:

$$\eta_k(\mathbf{S}'_k) = \frac{(\lambda_k/\mu)^{\mathbf{S}'_k+1}}{1-(\lambda_k/\mu)} \tag{3.13}$$

Based on the numerical experiments presented in Section 3.4, proposed approximation method is highly accurate when the inventory system constitutes only two bases. However, as the number of bases increases, the assumption of no part failure from other bases becomes more and more unrealistic which leads to poor approximations. As the number of bases increases, the time to revisit state $\mathbf{x}$ increases, and hence the expected cost incurred during this period at base k approaches to the average cost at that base ($\eta_k(\mathbf{S}'_k)$). The proposed approximate stock allocation policy is an index policy. If the current state is $\mathbf{x}$, index policy $\tilde{\Phi}(\mathbf{x})$ selects the base that maximizes the following expression:

$$\tilde{\Phi}(\mathbf{x}) = \arg\max_{k\in\mathsf{K}}\{\frac{\eta_k(\mathbf{S}'_k)}{\lambda_k+(K-2)\mu}\} \tag{3.14}$$

The proposed stock allocation policy considers the change in relative value function for two bases. It allocates the part to the base which realizes the highest reduction in backorder risk (measured by relative value function) if the part is allocated to that base. As the number of bases increase, it relies more on the average backorder cost of the bases rather than the relative value function. For each base in a K base model, the value of the expression presented in Equation

3.14 is obtained. The proposed stock allocation policy allocates the repaired part to the base with the largest value.

It was shown that finding the stock allocation stock allocation leads to finding the initial base-stock level and vice versa. Hence, based on the proposed stock allocation policy, optimal base-stock level can be approximated. Starting from zero on-hand inventory ($\mathbf{S}' = 0$), available service parts are allocated to the bases based on the proposed stock allocation policy until all the service parts are allocated. This results in the approximate base-stock level.

3.4 Numerical Study: Results and Discussion

To evaluate the performance of the proposed stock allocation policy (VF-based policy) compared with the optimal stock allocation policy and the myopic policy proposed by Perez and Zipkin [139], numerical experiments are conducted. Optimality gaps of the proposed stock allocation policy and the myopic rule are investigated and in order to obtain further insights on the performance of the heuristic policies, sensitivity analysis is performed.

To meet the first objective of the numerical experiment, a two-base model is considered and a test bed with a factorial design on four parameters is used. Value-iteration method is used to calculate the expected backorder cost for each policy. For the second objective, two experiments are designed. To investigate the impact of different parameters except the number of bases on the optimality gap, a two-base model is considered and the performance of each policy is evaluated by varying each of the parameters. The test bed and expected backorder cost calculation method are similar to the previous experiment. Since real world

Table 3.1: Input parameters for general evaluation of stock allocation policies

Parameters	Values
λ_2/λ_1	1/3,1/2,1,2,3
c_2/c_1	1,2,3
ρ	0.8,0.9
Total Stock	8,12

service parts supply chain constitute large number of bases, another experiment is carried out. The number of bases is varied from 3 to 10 and a simulation study is conducted to compare the performance of the proposed policy with the myopic policy. In all of the experiments, for the optimal policy, extensive solution space search is used to find the optimal base-stock level, while for the proposed stock allocation policy and the myopic policy, the base-stock calculation method described in Section 3.3.2 is used.

To investigate the optimality gap of the heuristic stock allocation methods, a test bed is used based on factorial design. Optimality gap for a particular stock allocation policy (f) is calculated as follows:

$$\Delta_f = \frac{100(\eta_f - \eta^*)}{\eta^*} \tag{3.15}$$

Parameter values used in the numerical study are shown in Table 3.1.

Combining all the parameters 52 problem instances are defined. Numerical results are summarized in Table A.1. Based on the results, following observations are made:

- Using the VF-based stock allocation rule resulted in 0.141 % optimality gap, while this number for the myopic policy was 0.644 %.

- Out of 52 cases, in 33 cases, base-stock level obtained by VF-based policy were same as the optimal base-stock level, while this number for myopic policy was 24.
- Out of 12 cases with highest optimality gaps of two approximate policies, 11 instances of VF-based policy and 8 cases of the myopic policy had different base-stock levels.
- As imbalance in the demand rates and backorder costs of the two bases, increased, optimality gap of the myopic policy increased, while the increase in the optimality gap of the VF-based policy was insignificant.

Sensitivity Analysis In addition to the general evaluation of the performance of the proposed stock allocation policy, sensitivity analysis of the parameters is performed. In all of the problem sets, it is assumed that $c_1 \leq c_2$. Followings are the results of the sensitivity analysis:

- **Demand Ratio:** Demand rates are varied between 1 to 10, resulting in a demand ratio of 0.1 to 10. For each case, cost is varied as $c = \{1, 2, 3\}$, while ρ is 0.8 and total initial stock is set to 10. As the demand ration increases, optimality gap of the VF-based policy decreases. Optimality gaps of the myopic policy are at the lowest levels, when the demand ratio is close to 1 and as the imbalance of demand rates increases, optimality gap increases. This fact verified earlier observation in the previous experiment (Figure 3.5).
- **Backorder Cost:** Backorder costs are set to $c_1 = 1, c_2 = \{1, 2, ..., 10\}$. Demand rates are varied between $\{1, 2, 3\}$ and ρ and total initial stock are

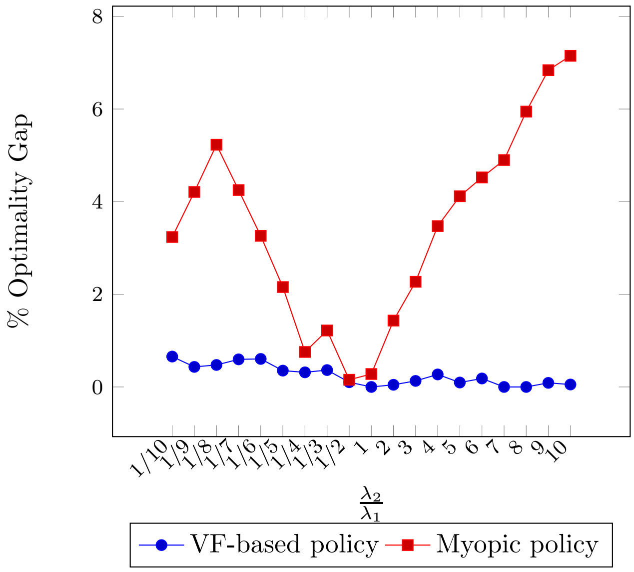


Figure 3.5: Demand sensitivity analysis

the same as the demand ratio sensitivity analysis. Based on the numerical results, as c_2/c_1 increases, optimality gaps of the VF-based policy has a slight increase with the slope of 0.005, while this number is 0.262 for the myopic stock allocation policy (Figure 3.6).

- **Repair Server Utilization:** Repair server utilization is set to be $\rho = \{0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$, while $c_1 = 1, c_2 = \{1, 2, 3\}$ and the demand rates are varied between $\{1, 2, 3\}$. In all of the cases, total initial stock is 10. As server utilization increases, the difference between optimality gaps of two heuristic policies decreases. Optimality gap of the VF-based policy reaches its minimum in $\rho = 0.8$, while for the myopic policy minimum optimality gap is achieved in $\rho = 0.9$ (Figure 3.7).

- **Investment Budget:** Initial stocks are varied between $\{6, 10, 14, 18, 22, 26\}$, while backorder cost and demand rates are the same as the previous exper-

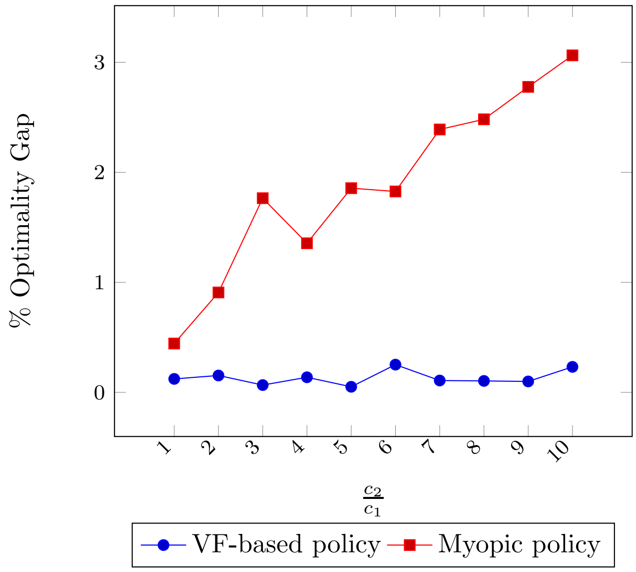


Figure 3.6: Backorder cost sensitivity analysis

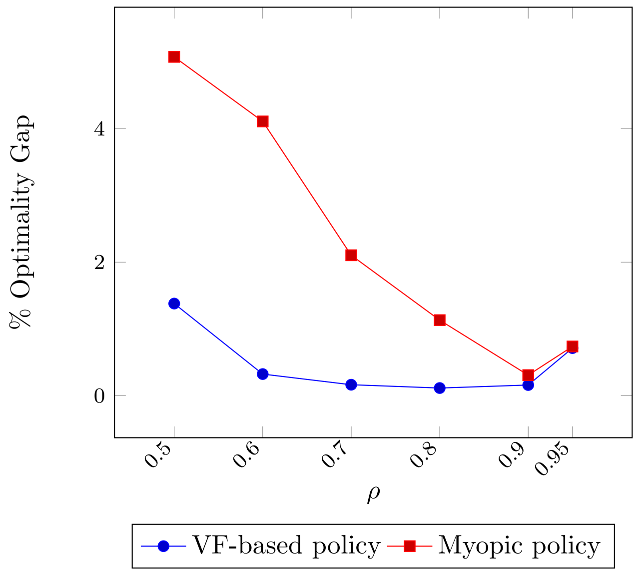


Figure 3.7: Repair server utilization sensitivity analysis

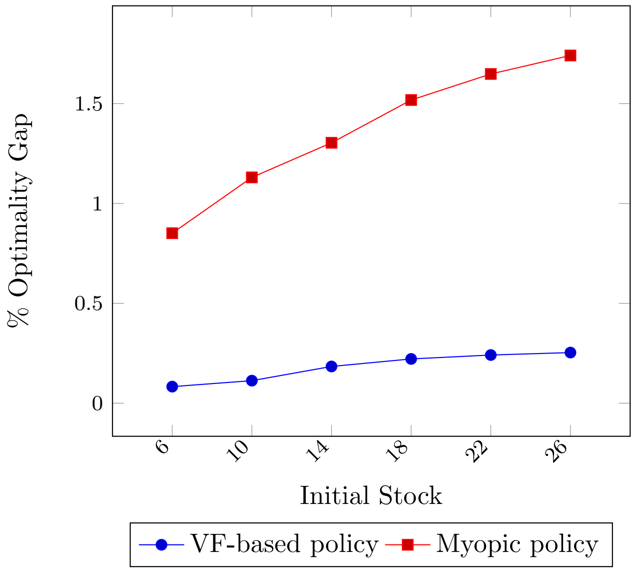
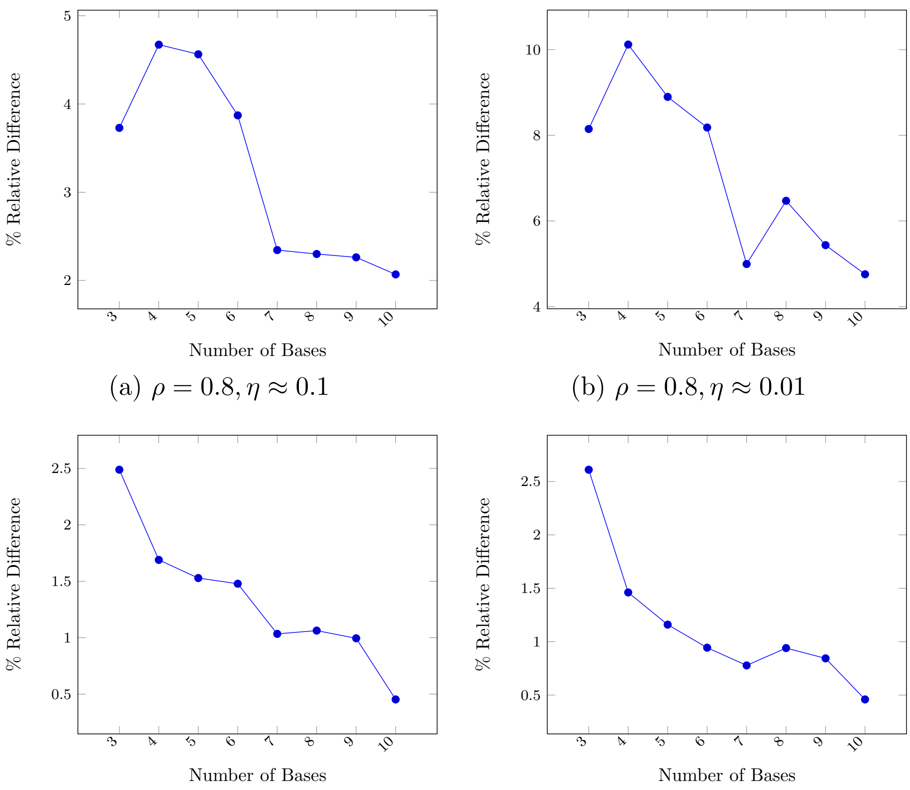


Figure 3.8: Initial budget sensitivity analysis

iment. Repair server utilization is set to be 0.8. It was observed that as the base-stock levels increases, optimality gaps of both policies increases, too. However, the amount of increase in VF-based policy is lower than the amount of increase in myopic policy. As the total base-stock level increases, the region in which both bases have stocks becomes larger. This region is the only region, where the stock allocation policy is used and as it becomes larger, it is reasonable to expect higher optimality gaps. (Figure 3.8).

- **Number of Bases:** To compare the proposed stock allocation policy with the myopic policy for higher number of bases, a simulation study is conducted. Inventory systems of 3 to 10 bases are considered. For each number of base, two cases with server utilization of 0.8 and 0.9 are investigated. For each case, the initial base-stock level is set to have the expected backorder cost close to 0.1 and 0.01. The combination of server utilization and initial base-stock level results in 4 combinations. For each combination, 100 problem sets are created with random demand rates and backorder costs. Demand rates are drawn from uniform distribution between $(0, K]$. It is assumed that there are three classes of customers with backorder costs of $\{1, 2, 3\}$ and each base is assigned to one of the classes, randomly. For each problem set, the proposed stock allocation policy and myopic policy are simulated for 10 replications of 100000 time steps. A warm-up period of 1000 time steps is also considered. Figure 3.9 illustrate the simulation result. Relative difference in each case is calculated by average backorder cost of 100 random problem sets for two policies.



(a) $\rho = 0.8, \eta \approx 0.1$

(b) $\rho = 0.8, \eta \approx 0.01$

(c) $\rho = 0.9, \eta \approx 0.1$

(d) $\rho = 0.9, \eta \approx 0.01$

Figure 3.9: Number of bases sensitivity analysis

As the number of bases increases, the relative difference between VF-based policy and myopic policy decreases. Increasing the number of bases has little impact on the total initial stock. Hence, as the number of base increases, the number of stocks allocated to each base decreases, significantly. Subsequently, as it was observed in initial budget sensitivity analysis, the performance difference of two policies decreases. Additionally, when the server utilization is 0.8, average backorder cost plays a key role in the relative difference of two policies, for the expected backorder cost of around 0.01, it is almost twice as the cases with expected backorder cost of around 0.1. When the expected backorder cost is low, higher number of parts are allocated to the bases and thus the relative difference of two policies increase. However, when the server utilization is 0.9, the effect of average backorder cost is insignificant.

3.5 Summary

A continuous-review inventory model with stock allocation decision made by the repair facility was considered. The problem was modeled as a Markov Decision Process and it was shown that the stock allocation switching curve passes the optimal initial base-stock level and the solution to the optimal initial base-stock level can be found by finding the optimal stock allocation policy and vice versa. A stock allocation method was proposed based on the combination of the value function and average backorder cost of a single base. For an inventory system with two bases, it is completely based on the relative value function of a single base and as the number of bases increase, it relies more on the average backorder cost of a single base. Extensive numerical experiments were carried out to evaluate the performance of the proposed policy. Based on the numerical results, proposed policy fell below the optimal policy only by only 0.141 % on average cost.

Chapter 4

Stock Reallocation of Service Parts among the Bases

4.1 Introduction

In the previous chapter, a stock allocation policy based on the value function of a single base was presented. Based on the numerical results, incorporating such flexibility in the stock allocation process in the repair facility leads to cost savings due to the decreased number of backorders. In this chapter, a stock reallocation policy is developed. Stock reallocation policy involves the movement of spare parts among bases in order to decrease the backorder risk. As mentioned in Chapter 2 stock reallocation has been addressed as unidirectional and bidirectional lateral transshipments as well as inventory rationing policies in the literature. The method explained in this chapter addresses both lateral transshipment and inventory rationing.

Following the stock allocation switching curves, in this chapter stock reallocation regions are found. During the inventory operation, when the inventory state enters these regions, a stock reallocation is initiated in order to exit this region. The method developed in this chapter is built upon the stock allocation model developed in the previous chapter and it is assumed that during the stock allocation policy synthesize, stock reallocation of service parts are not considered.

The problem of the optimal stock reallocation is modeled as a Markov Decision Process (MDP) and the dynamic program of the problem is provided. Afterwards, the relative value function of the respective stock reallocation policy is approximated. The performance of the proposed method is evaluated by numerical experiments, where it is compared with the optimal policy.

The remainder of this chapter is organized as follows: The problem under study is described and model formulation is presented in Section 4.2. In Section 4.3, the problem of optimal stock reallocation is addressed and the approximation method is described. Numerical experiments are conducted and the results are discussed in Section 4.4. Section 4.5 concludes with a summary of the contributions and findings of this chapter.

4.2 Problem Description and Model Formulation

As mentioned earlier, the stock reallocation model developed in this chapter has built upon the model developed in the previous chapter. However, some modifications are needed in order to address the stock reallocation problem. In this section, modifications made on the inventory system configurations and assumptions are explained and the problem formulation is presented.

4.2.1 System Description

Consider the repairable service parts inventory system presented in the previous chapter. At the beginning of operation, initial stocks are provisioned and allocated to each base. During the operation, when a part is repaired, repair facility allocates the ready-to-use part to one of the bases based on the stock allocation policy proposed in the previous chapter. When the inventory level at one of the bases is low, it is reasonable to reallocate a part from a base to another in order to decrease the backorder risk. This approach seems to be reasonable considering the fact that usually repairable service parts are used in systems where uptime is crucial and reallocation cost is significantly lower than backorder costs. The objective is to find the time a reallocation is needed and sender and receiver bases in order to minimize the expected backorder and reallocation costs. Figure 4.1 illustrates a graphical representation of the inventory network and reallocation of service parts within the resupply network.

4.2.2 Notations and Assumptions

All of the assumptions made in Chapter 3 applies to the model developed in this chapter. Additionally, it is assumed that service parts are transshipped between the bases with exponential time at the rate of τ. The assumption of exponential transshipment time is required in order to be able to model the problem using MDP and it captures the emergency of proactively transshipment of parts among bases within a limited time. For the sake of simplicity, it is assumed that τ is

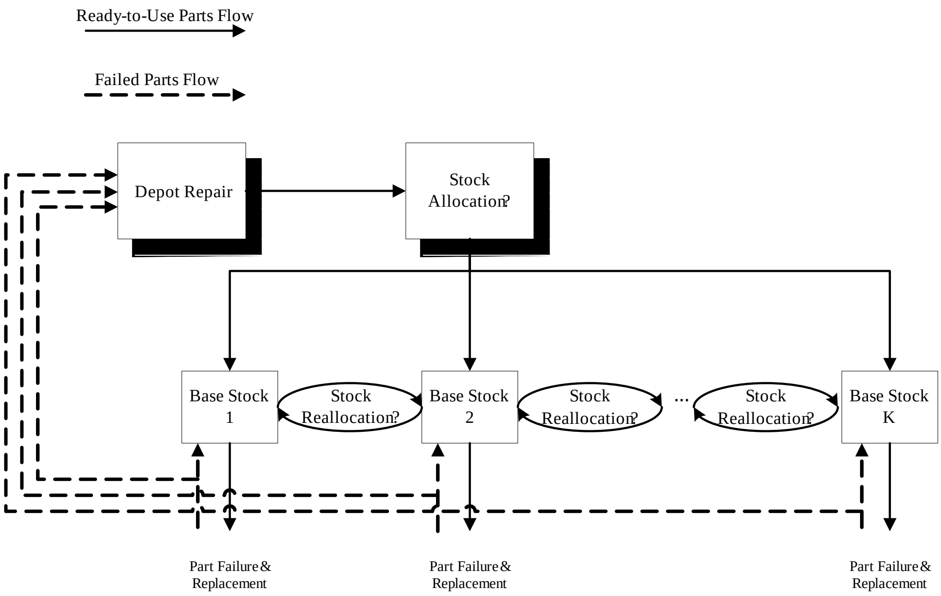


Figure 4.1: Service parts inventory system with stock reallocation

identical for all of the inventory holding locations which are at different airports. If the inventory holding location are at the same airport, transshipment time is assumed to be zero.

4.2.3 MDP Formulation

To formulate this inventory system, same as the previous chapter the infinite time horizon model transformed into a discrete-time Markov Decision Process is considered and the modeling approach presented in Meyn [135] is used. Let $\mathbf{R}(t)$ denote the stock reallocation vector at time t, where $\mathbf{R}_k(t) = \{-1, 0, 1\}$, $\sum_{k=1}^{K} |\mathbf{R}_k(t)| = \{0, 2\}$ and $\mathbf{R}_k(t) \leq 0$ whenever $\mathbf{X}_k(t) \geq \mathbf{S}_k$. The queue length process at time step t at base k ($\mathbf{X}_k(t)$ increases by 1 with probability Λ_k, decreases by 1 with probability $\mathbf{F}_k(t-1)M$ and depending on the stock reallocation policy, it changes by 1 with the probability $\mathbf{R}_k(t-1)T$, where probabilities Λ_k,

M and T can be defined as follows:

$$\Lambda_k = \frac{\lambda_k}{\tau + \mu + \sum_{m=1}^{K} \lambda_m} \quad (4.1)$$

$$M = \frac{\mu}{\tau + \mu + \sum_{m=1}^{K} \lambda_m} \quad (4.2)$$

$$T = \frac{\tau}{\tau + \mu + \sum_{m=1}^{K} \lambda_m} \quad (4.3)$$

The repair and allocation process at time step t can be defined by the following recursion:

$$\mathbf{X}(t) = \mathbf{X}(t-1) - B(t)\mathbf{F}(t-1) + C(t)\mathbf{R}(t-1) + \mathbf{A}(t) \quad (4.4)$$

where the i.i.d process $(\mathbf{A}_k(1 \leq k \leq K), B, C)$ has the marginal distribution defined by:

$$\mathsf{P}\{(\mathbf{A}_1(t), \mathbf{A}_2(t), ..., \mathbf{A}_K(t), B(t), C(t)) = \mathbf{e}_i\} = \mathbf{p}_i, \quad (4.5)$$

with $\mathbf{p}_i = \Lambda_i$ for $i = 1, 2, ..., K$, $\mathbf{p}_{K+1} = M$ and $\mathbf{p}_{K+2} = T$. In this case, $(\mathbf{X}, \mathbf{R})$ is a Markov Decision Process, where the stock reallocation decision $\mathbf{R}(t)$ is defined based on the number of parts in the repair facility from each base.

Let π denote the stock reallocation cost from a base to another. Same as τ, stock reallocation cost is assumed to be identical between the inventory holding locations at different airports. For the inventory holding locations at the same airport, it is assumed to be zero. Let $\mathbf{G}^r(\mathbf{x})$ denote the stock reallocation vector when the inventory state is $\mathbf{x}$ and reallocation of parts are between two airports. If the reallocation is between two bases at the same airport $\mathbf{G}^r(\mathbf{x}) = 0$. Stock reallocation cost when the queue length is $\mathbf{x}$ under a particular stock reallocation policy r is:

$$\pi_r(\mathbf{x}) = T\pi\mathbf{G}^r(\mathbf{x}). \tag{4.6}$$

Let $\mathbf{c}_k$ denote the backorder cost at base k per unit per time. Assuming the current state is $\mathbf{x} = \mathbf{X}(t)$, the stage cost at time t is:

$$C_r(\mathbf{x}) = \pi_r(x) + \sum_{k=1}^{K} \mathbf{c}_k[\mathbf{x}_k - \mathbf{S}_k]^+. \tag{4.7}$$

As mentioned earlier, when service parts are reallocated between two bases at the same airport, the reallocation is instant with zero cost. If the current state is $\mathbf{x}$ and a part fails at base $k1$, the state changes to $\mathbf{x} + \mathbf{e}_{k1}$. If it is optimal to reallocate a part from base $k2$ to base $k1$, the inventory state changes $\mathbf{x} + \mathbf{e}_{k2}$, instantly. In this case, instead of modeling the reallocation of part from the base $k2$ to the base $k1$ when the inventory state is $\mathbf{x} + \mathbf{e}_{k1}$, one can consider a model in which when the state is $\mathbf{x}$, demand from base $k1$ and $k2$ is fulfilled by the base $k2$. Let $\mathbf{e}'$ denote such demand fulfillment vector. If there is no stock reallocation or the reallocated part is to be transshipped between two airports, $\mathbf{e}'_k = \mathbf{e}_k$. However, when there is a reallocation of part between bases $k1$ and $k2$ as explained, $\mathbf{e}'_{k1} = \mathbf{e}_{k2}$, while $\mathbf{e}'_k = \mathbf{e}_k$ for all other bases.

4.3 Stock Reallocation Policy

Let $h^*(\mathbf{x})$ and $\eta^*(\mathbf{S})$ denote the relative value function for the stock reallocation dynamic program and the steady-state average cost of the system under the optimal stock reallocation policy, respectively. The average cost is independent of the initial state $\mathbf{x}$ and the dynamic programming optimality equation [135] can be defined as:

$$\eta^*(\mathbf{S})+h^*(\mathbf{x}) = \min_{r\in \mathsf{R}(\mathbf{x})}[C_r(\mathbf{x})+(\sum_{k=1}^{K}\Lambda_k h^*(\mathbf{x}+\mathbf{e}'_k)+Mh^*(\mathbf{x}-\mathbf{e}_f)+Th^*(\mathbf{x}+G_r(\mathbf{x}))]. \tag{4.8}$$

Without loss of generality it can be assumed that the state of origin is 0. Let $\Phi'(x)$ denote the stock reallocation policy when the current state is $\mathbf{x}$. Optimal stock reallocation policy can be defined based on the optimal relative value function as follows:

$$\Phi'^*(x) = \arg\min_{r\in \mathsf{R}(\mathbf{x})}[C_r(\mathbf{x}) + (\sum_{k=1}^{K}\Lambda_k h^*(\mathbf{x}+\mathbf{e}'_k) + Mh^*(\mathbf{x}-\mathbf{e}_f) + Th^*(\mathbf{x}+\mathbf{G}^r(\mathbf{x}))]. \tag{4.9}$$

Figure 4.2 illustrates the optimal stock reallocation regions for the example shown in Figure 3.3, where stock reallocation cost and rate from each base to another are set to be 0.01 and 30, respectively. As can be seen, stock reallocation regions prevent the queue length process from entering to the regions of imbalance, that is having more of inventory in one of the bases and less inventory in the other. Consider the case, where the inventory level of this inventory system is $[8,3]$ or the inventory state is $\mathbf{x} = [8,5]$. If the next event is a part failure at base 2, its inventory level drops to 2 and the inventory state becomes $\mathbf{x} = [8,6]$. Since this inventory state is inside the region of imbalance or in other words $\mathbf{G}^r([8,6]) = [-1,1]$, a ready-to-use part is transshipped from base 1 to base 2 in order to decrease the backorder risk at base 2. If no part fails and no repair job finishes until the transshipment is done, the inventory level changes to $[7,3]$.

Calculation of h^* is intractable for large number of bases. In this section, h^* is

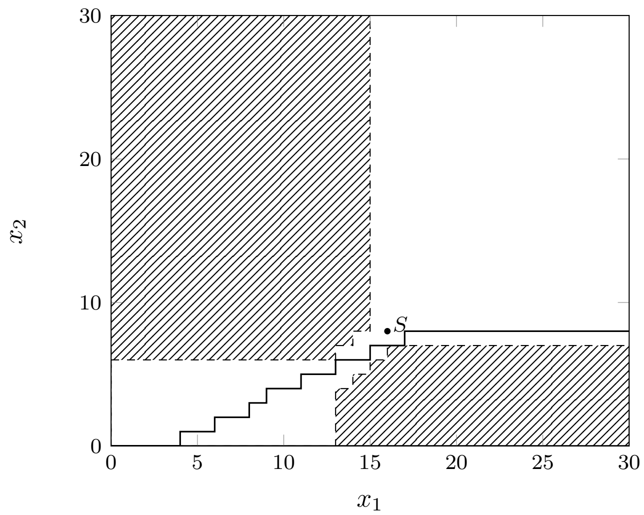


Figure 4.2: Optimal stock reallocation policy and regions of imbalance

approximated based on the so-called state-space collapse phenomenon. Figure 4.3 demonstrates a sample path for the example shown in Figure 3.3. Additionally, Figure 4.4 shows the probability distribution of the queue length based on a simulation of 10^7 steps for 10 replications for the same example. As can be seen, sample path and the steady-state probability distribution of the queue lengths are highly concentrated around the stock allocation switching curve.

State-space collapse phenomena has been studied extensively by the heavy-traffic literature on stochastic networks, where it is mainly used to construct a parameterized family of networks, where the n-dimensional process is approximated by a controlled Brownian motion. Meyn [135] uses the state space collapse for workload relaxation of complex networks. In a workload relaxation, a network of large size is approximated by a queue with weighted sum of network-wide queueing costs. Stolyar [140] considers a general switch model, similar to the stock allocation model considered in the previous chapter. He considers a heavy traffic regime and proves that under the Max-Weight scheduling and resource

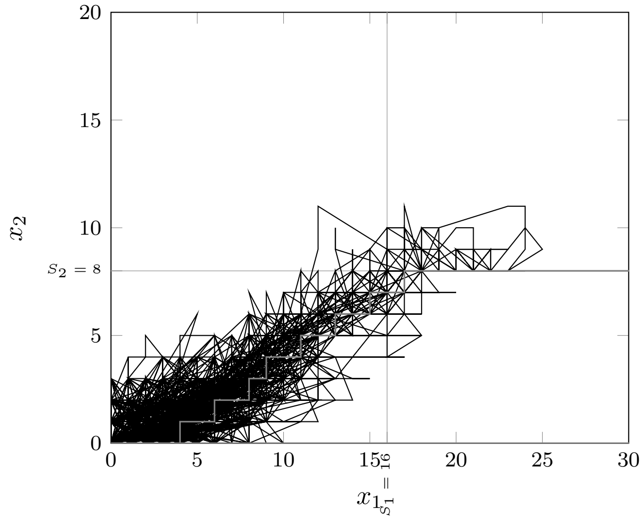


Figure 4.3: A sample path of a two-base model with stock allocation

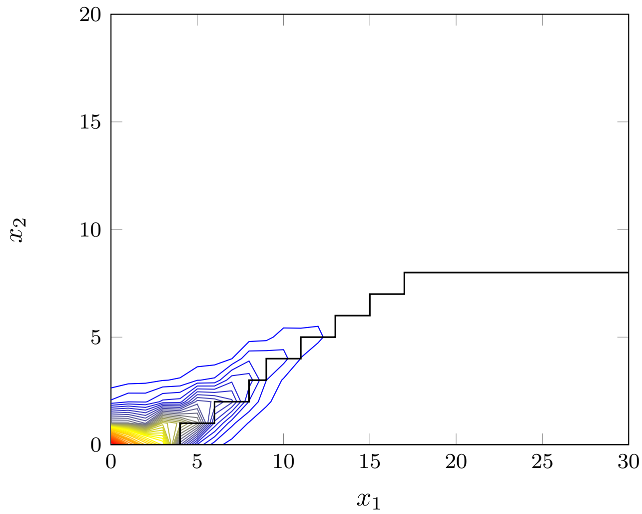


Figure 4.4: Probability distribution contour of a two-base model with stock allocation

pooling condition, the queue length process has the state space collapse property. Max-Weight scheduling discipline chooses a decision which maximizes the rate of decrease in a predetermined queueing cost function. Bramson [141] applies the state space collapse phenomena to two multi-class queueing networks: first-in-first-out queueing networks of Kelly type and head-of-the-line proportional processor sharing queueing networks. He proves that the state space collapse holds for more general queueing networks, provided the solutions of their fluid model equations converge.

Fluid network model is a macroscopic crude representation of its stochastic counter-part, where the uncertainty property of the network is removed and jobs inside the queues are considered to be continuous [142, 135]. In other words, the fluid network model represents the mean behavior of its stochastic counter-part. If the inventory model considered in Figure 4.3 is replaced with its fluid counter-part, starting from any state x, the sample path moves towards the stock allocation switching curve and evolves on the switching curve afterwards. This phenomena is called state space collapse. Stock allocation switching curve can be considered as the one-dimensional state space of a single base model which approximates the behavior of the K base inventory model with K dimensional state space.

Research into the fluid model of the inventory system considered in this chapter is out of the scope of this study. However, in order to validate the use of state space collapse phenomena, some properties of the fluid model for the inventory system considered here is necessary. One can note that in the region $\mathcal{R}_s$ (Figure 3.3), fluid value function is zero as no cost is incurred in this region. Since the optimal stock allocation policy in region $\mathcal{R}_b$ (Figure 3.3) is to send the part to the base with highest backorder cost, one can write down the rates of change in the fluid model on the boundary states and the states inside $\mathcal{R}_b$. Hence, fluid value function in region $\mathcal{R}_b$ can be found by solving the respective system of partial differential equations. The fluid value function $J^*(\mathbf{x})$ can be expressed as as follows:

$$J^*(\mathbf{x}) = 1/2([\mathbf{x} - \mathbf{S}]^+ \mathbf{D} [\mathbf{x} - \mathbf{S}]^{+T}), \tag{4.10}$$

where D is defined as follows:

$$\begin{aligned} \mathbf{D}_{i,j} &= \mathbf{D}_{k,k}, \qquad k = \min\{i, j\}, \\ \mathbf{D}_{i,i} &= \frac{\mathbf{c}_i + \sum_{k=1}^{i-1} \mathbf{D}_{k,k} \lambda_k}{\mu - \sum_{k=i}^{K} \lambda_k}. \end{aligned} \tag{4.11}$$

Hence, state space collapse holds for the service parts inventory model with stock allocation policy presented in the previous chapter.

4.3.1 Approximating h^*

Based on the state space collapse phenomena, the queue length process of an inventory system with K bases can be approximated using an aggregate queue length process, where the demand rate for the aggregate queue is the sum of demand rates at K bases. Hence, in order to approximate the relative value function of the inventory system with K bases, the relative value function of a single base model is derived. Consider a service parts inventory system with a base and a repair facility with single non-idling server. Let ρ and c denote the server utilization and unit backorder cost of this single queue system. Average cost of the single base model is as follows:

$$\eta^* = c\frac{\rho^{S+1}}{1-\rho} \tag{4.12}$$

The relative value function for the single base model is a piecewise function, $h_s^*(x)$ when there is on-hand inventory and $h_b^*(x)$ when there is no inventory on-hand or there is backorder. Relative value function for a single base model for $0 \le x \le S$ is defined by $h_s^{1*}(x)$:

$$h_s^*(x) = \frac{\eta^*}{\lambda}\sum_{k=0}^{x}(x-k)\frac{1}{\rho^k} \tag{4.13}$$

Relative value function for a single base model when $x > S$ is defined by $h_b^*(x)$

:

$$h_b^*(x) = \frac{\eta^*(\sum_{k=0}^{x}\frac{x-k}{\rho^k})}{\lambda} - c\frac{\mu(\sum_{k=1}^{x-S}(\frac{1}{\rho^{k-1}} - k(1-\rho))) - (x-S)\lambda}{(\mu-\lambda)^2} \tag{4.14}$$

The relative value function for a single base model can be used as the approximate relative value function of the K base model when the current state $\mathbf{x}$ is on the stock allocation switching curve. The demand rate of the aggregate model is the sum of demand rates at the bases and its backorder cost is $\min_{k\in K}\{c_k\}$. To approximate the relative value function, when the state is not on the stock allocation switching curve, another single base model is considered, where service rate can be either μ or $\mu+\tau$, depending on the state. Let l^* denote the state where for $x < l^*$, service rate of the queue is μ and for $x \ge l^*$, the service rate is $\mu+\tau$ and a transition cost of $\pi\frac{\tau}{\tau+\mu}$ is incurred. In this case, l^* minimizes the following average cost function:

$$\eta(S,l^*) = \frac{c(\mu-\lambda)(\frac{\lambda}{\mu})^{l^*}\lambda\,(\frac{\lambda}{\tau+\mu})^{S-l^*}}{(\tau+\mu-\lambda)(\tau+\mu-\lambda-(\frac{\lambda}{\mu})^{l^*}\tau)} + \frac{\pi\tau(\mu-\lambda)(\frac{\lambda}{\mu})^{l^*}}{(\lambda+\mu+\tau)(\tau+\mu-\lambda-(\frac{\lambda}{\mu})^{l^*}\tau)} \tag{4.15}$$

The relative value function of the single base model when $x < l^*$ is the same as the previous single base model, provided that the time is rescaled so $\lambda+\mu+\tau = 1$. Relative value function for a single base model when $x \geq l^*$ is defined by $h_l^*(x)$:

$$h_l^*(x) = \frac{((\frac{\mu}{\lambda})^{l^*}-1)\eta\lambda(\frac{\mu+\tau}{\lambda})^x}{(\mu+\tau-\lambda)(\mu-\lambda)} + \frac{(\tau\,\lambda(\frac{\mu}{\lambda})^{l^*}+(\lambda-\mu)(\mu+\tau-\lambda)l^*-\tau\lambda)\eta}{(\mu-\lambda)^2(\mu+\tau-\lambda)} + \frac{(\eta-c\tau)(\lambda(\frac{\mu+\tau}{\lambda})^x+\mu+\tau-2\lambda)}{(\mu+\tau-\lambda)^2} + \frac{(c\tau-\eta)(x+1)}{\tau+\mu-\lambda}. \tag{4.16}$$

Let $\tilde{h}(\mathbf{x})$ denote the approximate relative value function. An approximation method based on the sum of the relative value function of series of single base models is proposed as follows:

$$\tilde{h}(\mathbf{x}) = h_s^*(\bar{\mathbf{x}}) + \sum_{k=1}^{K} h_l^*(\mathbf{x}_k) \tag{4.17}$$

where $\bar{\mathbf{x}} = \sum_{k=1}^{K}\mathbf{x}_k$

Approximate stock reallocation can be defined based on the approximate relative value function as follows:

$$\tilde{\Phi}'(\mathbf{x}) = \arg\min_{r\in\mathsf{R}(\mathbf{x})}[C_r(\mathbf{x})+(\sum_{k=1}^{K}\Lambda_k\tilde{h}(\mathbf{x}+\mathbf{e}'_k)+M\tilde{h}(\mathbf{x}-\mathbf{e}_f)+T\tilde{h}(\mathbf{x}+\mathbf{G}^r(\mathbf{x}))]. \tag{4.18}$$

Proposed stock reallocation policy approximates the backorder risks at the bases using the relative value function. It allocates a part from the base with minimum backorder risk to the base with maximum backorder risk, when the difference exceeds the reallocation cost. As can be seen, stock reallocation policy is defined based on the vector $\mathbf{r}$, which is the minimizer of the expression presented in Equation 4.18 and has the same properties as $\mathbf{R}(t)$ presented in Section 4.2.3. If $\mathbf{r} = 0$, it implies that no reallocation of service part from one base to another is needed. However, if $\mathbf{r}_{k1} = -1$ and $\mathbf{r}_{k2} = 1$, reallocation of a service part from the base $k1$ to the base $k2$ is initiated.

In the proposed approximation, the relative value function of a K base model is approximated by the sum of the relative value functions of the aggregate queue and each individual single queue. In other words, the K base inventory system is treated as $K + 1$ independent queues with a queue representing the aggregate behavior of the inventory system and K queues representing the behavior of each individual base. Based on the numerical results, it is accurate and applicable for the inventory systems with any number of bases.

4.4 Numerical Study: Results and Discussion

To evaluate the performance of the proposed stock reallocation policy compared with the optimal stock reallocation policy and investigate the possible cost reductions by utilizing the stock reallocation policy, numerical experiments are conducted. The optimality gap of the proposed stock reallocation policy is investigated and in order to obtain further insights on the performance of the approximate policy, sensitivity analysis is carried out.

To meet the first objective of the numerical experiment, a two-base model

Table 4.1: Input parameters for general evaluation of stock reallocation policies

Parameters	**Values**
λ_2/λ_1	1/3,1/2,1,2,3
c_2/c_1	1,2,3
ρ	0.8
Total Stock	8,12
π	0.01,0.05
τ	5μ,10μ

is considered and a test bed with a factorial design on six parameters is used. Average cost is used as the performance measure for both policies and was calculated by the value iteration method. Stock allocation policy and initial base-stock levels is obtained using the method described in Section 3.3.2. For the second objective, two experiments are designed. To investigate the impact of different parameters except the number of bases on the optimality gap, a two-base model is considered and the performance of each policy is evaluated while varying each of the parameters. The test bed and average cost calculation method are similar to the previous experiment. Since real world service parts supply chain constitute large number of bases, another experiment is carried out, in which the number of bases is varied from 3 to 10 and a simulation study is conducted to compare the performance of the proposed stock reallocation policy with the inventory system without stock reallocation policy.

To investigate the optimality gap of the approximate stock reallocation policy, a test bed is used based on factorial design. Parameter values used in the numerical study are shown in Table 4.1.

Combining all the parameters 104 problem instances are defined. Numerical results is summarized in Tables A.2 and A.3. It was observed that introducing the stock reallocation capability to the inventory system resulted in 4.871% reduction in cost. Additionally, Approximate stock reallocation policy performed very close to the optimal policy with maximum and average optimality gaps of 0.682% and 0.090%, respectively.

Sensitivity Analysis In addition to the general evaluation of the performance of the proposed stock reallocation policy, sensitivity analysis of the parameters is performed. In all of the problem sets, it is assumed that $c_1 \leq c_2$. Followings are the results of the sensitivity analysis:

- **Demand Ratio:** Demand rates are varied between 1 to 10, resulting in a demand ratio of 0.1 to 10. For each case, backorder cost is varied as $c = \{1, 2, 3\}$, while ρ is 0.8, total initial stock is set to 10, $\tau = 5\mu$ and $\pi = 0.01$. As the demand ratio increases, optimality gap of the approximate stock reallocation policy increases and the amount of cost reduction decreases (Figure 4.5). Based on the structure of the stock reallocation policies,independent stock allocation policy assumption is the main source of these trends. Since $c_2/c_1 \geq 1$, as the demand ratio increases, region of imbalance for the base with higher backorder cost moves towards the stock allocation switching curve and is blocked when it reaches to this curve.

- **Backorder Cost:** Backorder costs are set to $c_1 = 1$ and $c_2 = \{1, 2, ..., 10\}$. Demand rates are varied between $\{1, 2, 3\}$ and other parameters are the same as the demand ratio sensitivity analysis. Based on the numerical results, no particular trends are observed in terms of the optimality gaps and cost reductions (Figure 4.6).

- **Repair Server Utilization:** Repair server utilization is set to be $\rho =$

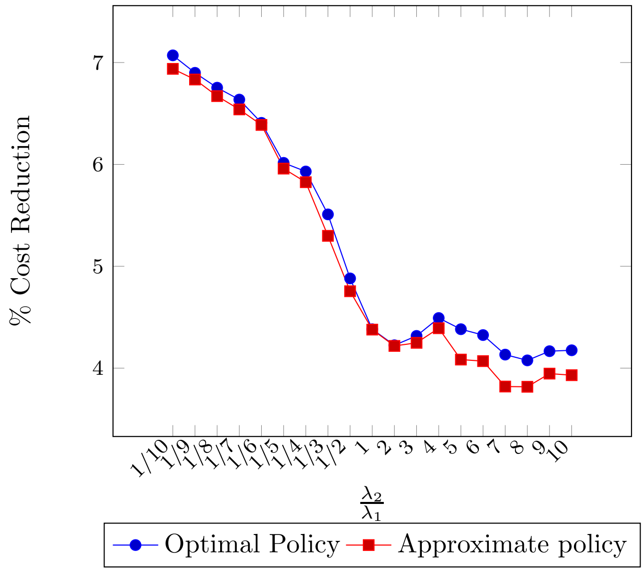


Figure 4.5: Demand sensitivity analysis

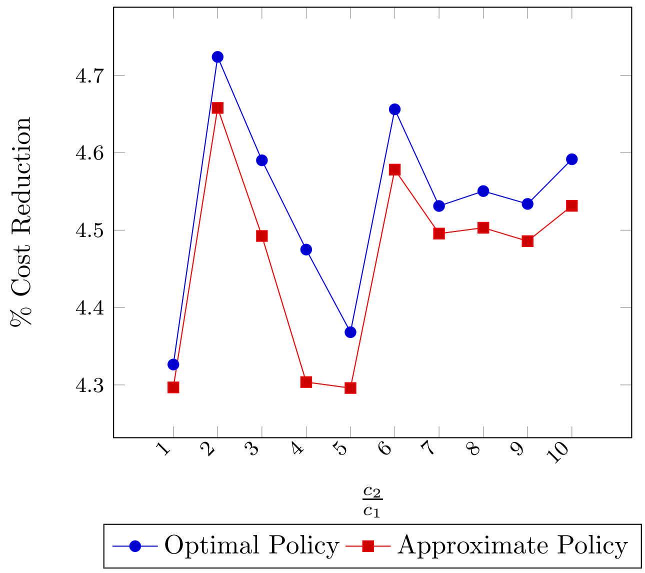


Figure 4.6: Backorder cost sensitivity analysis

$\{0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$, while $c_1 = 1, c_2 = \{1, 2, 3\}$ and all other parameters are the same as the previous experiment. It was observed that as server utilization increases, the amount of cost reduction decreases and reaches its minimum in $\rho = 0.9$. Stock reallocation is defined in regions $\mathcal{R}_s$, $\mathcal{R}_1$ and $\mathcal{R}_2$ (Figure 3.3). As the server utilization increases, incurring backorder cost in regions $\mathcal{R}_b$, $\mathcal{R}_1$ and $\mathcal{R}_2$ (Figure 3.3) increases. When the server utilization increases from 0.5 to 0.9, incurring backorder cost in region $\mathcal{R}_b$ increases and since the stock reallocation policy has no influence in this region, the amount of cost reduction decreases. However, when the server utilization goes from 0.9 to 0.95, although the incurring backorder cost in region $\mathcal{R}_b$ increases, its not significant in comparison with the backorder cost increase in regions $\mathcal{R}_1$ and $\mathcal{R}_2$, where utilizing the stock reallocation policy results in cost reduction. Hence, there is a slight increase in the amount of cost reduction as the server utilization goes from 0.9 to 0.95. Optimality gaps are independent of the server utilization (Figure 4.7).

- **Investment Budget:** Initial stock is varied between $\{6, 10, 14, 18, 22, 26\}$, while repair server utilization is set to be 0.8 and all other parameters are the same as the previous experiment. It was observed that as the base-stock level increases, the amount of cost reduction increases, too. As the total initial stock increases, the number of parts available for pooling between bases increases, too. However, as the initial stock increases, the rate of increase in cost savings decreases. No particular trend is observed in terms of the optimality gaps, when the initial stock is changed (Figure 4.8).

- **Stock Reallocation Cost:** Stock reallocation cost is varied between

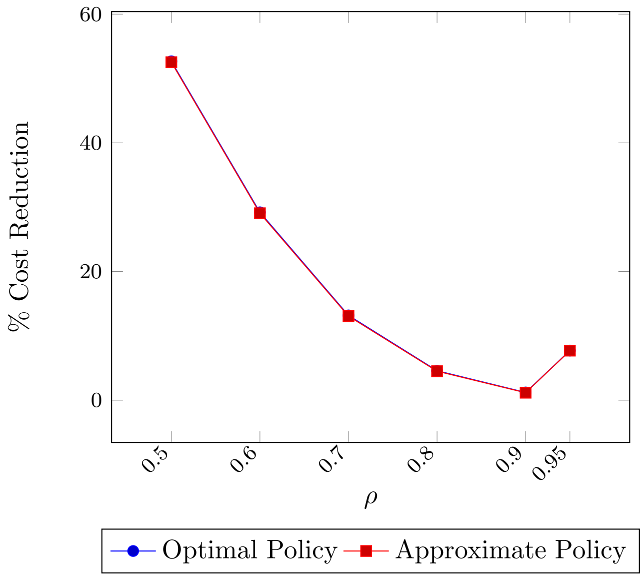


Figure 4.7: Repair server utilization sensitivity analysis

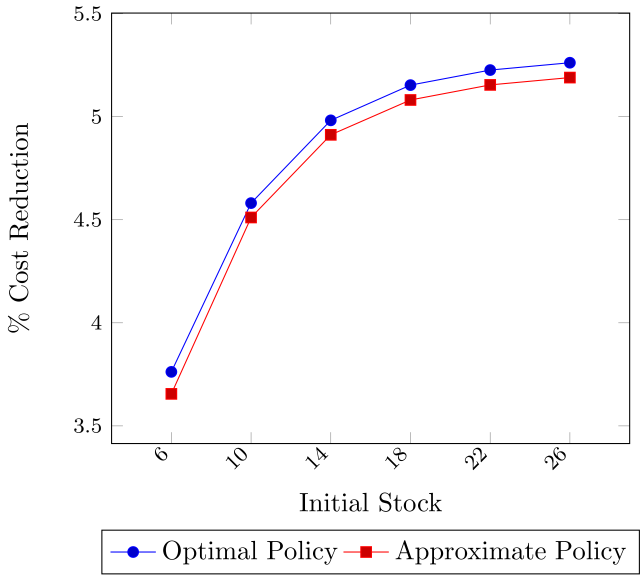


Figure 4.8: Initial budget sensitivity analysis

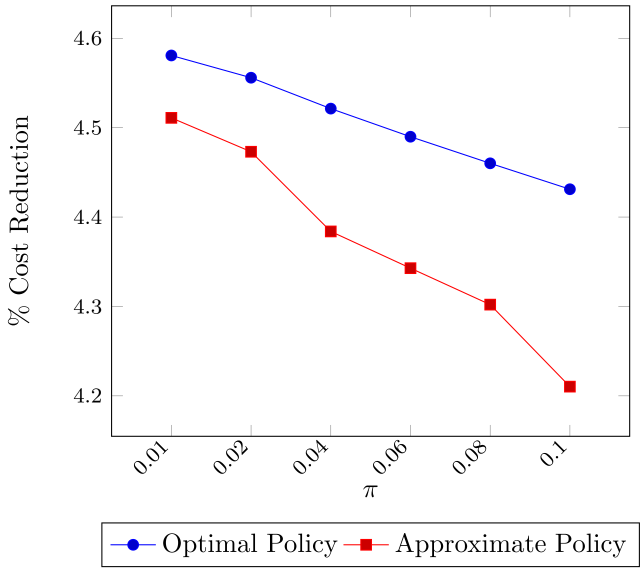


Figure 4.9: Stock reallocation cost sensitivity analysis

$\{0.01, 0.02, 0.04, 0.06, 0.08, 1\}$, while initial total stock is set to be 10 and all other parameters are the same as the previous experiment. It was observed that as the stock reallocation cost increases, the amount of cost reduction decreases. This is reasonable considering the fact that higher stock reallocation costs results in lower number of stock reallocation instances. Additionally, it is observed that as the stock reallocation cost increases, accuracy of the approximate stock reallocation policy decreases (Figure 4.9).

- **Stock Reallocation Rate:** Stock reallocation rate is varied between $\{2.5\mu, 5\mu, 7.5\mu, 10\mu, 12.5\mu, 15\mu\}$, while stock reallocation cost is set to be 0.01 and all other parameters are same as the previous experiment. It was observed that as the rate of stock reallocation increases, the amount of cost reduction increases, too. This is mainly because of the fact that when the stock reallocation rate increases, parts are reallocated between

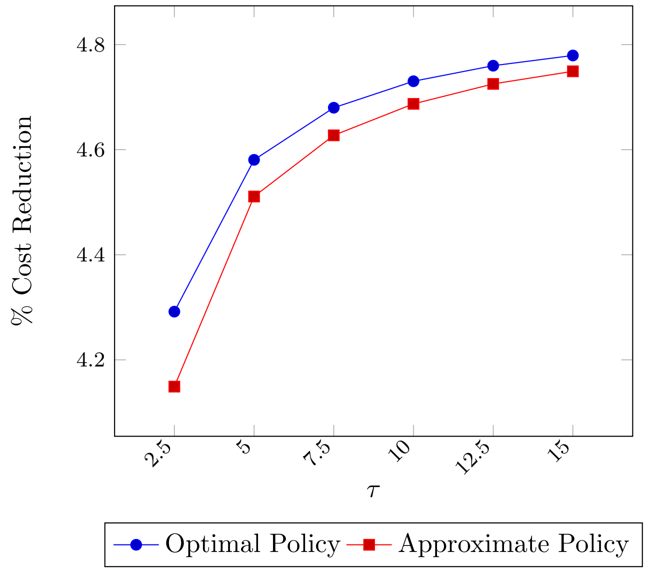
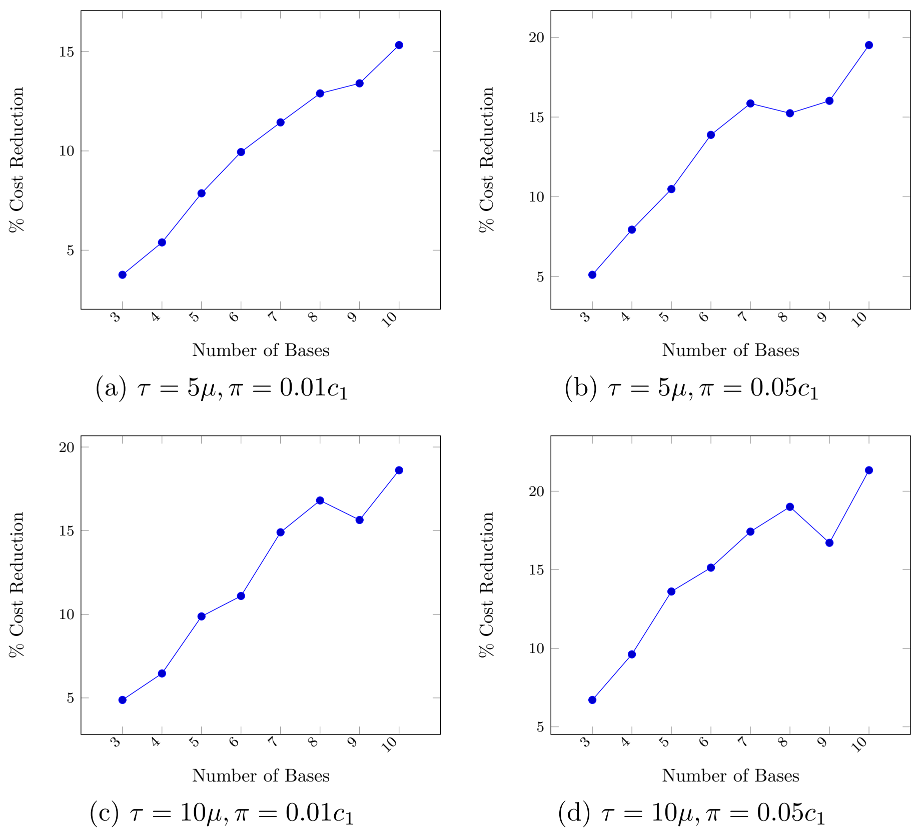


Figure 4.10: Stock reallocation rate sensitivity analysis

two bases swiftly and possibility of having backorders during this period decreases. Hence, introducing stock reallocation in such cases proves to be more effective. However, as the stock reallocation rate increases, the rate of increases in the amount of cost reduction decreases. Additionally, it was observed that approximate policy is more accurate for higher stock reallocation rates (Figure 4.10).

- **Number of Bases:** To investigate the amount of cost reductions from using proposed stock reallocation policy for higher number of bases, a simulation study is conducted. Inventory systems of 3 to 10 bases are considered. For each number of base, two cases with stock reallocation cost of $0.01c_1$ and $0.05c_1$ are considered. For each case, the stock reallocation rate was set to be 5μ and 10μ. The combination of stock reallocation cost and rate results in 4 combinations. The initial base-stock levels is set to have an expected backorder cost of around 0.1. For each combination, 50 problem sets are created with random demand rates and backorder costs. Demand rates are drawn from uniform distribution between $(0, K]$. It is assumed that there are three classes of customers with backorder costs of $\{1, 2, 3\}$ and each base is assigned to one of the classes, randomly. For each problem set, the proposed stock reallocation policy is simulated for 5 replications of 500000 time steps. A warm-up period of 500 time steps is also considered. Figure 4.11 illustrates the simulation result. Cost reduction in each case is calculated by the average cost of 50 random problem sets for two policies.



(a) $\tau = 5\mu, \pi = 0.01c_1$

(b) $\tau = 5\mu, \pi = 0.05c_1$

(c) $\tau = 10\mu, \pi = 0.01c_1$

(d) $\tau = 10\mu, \pi = 0.05c_1$

Figure 4.11: Number of bases sensitivity analysis

As can be seen, as the number of bases increases, the amount of cost reduction from using the stock reallocation policy increases, too. As the number of bases increase, there are more sources to deliver a part using stock reallocation to a base with low inventory level. Hence, implementing the stock reallocation policy in large resupply networks results in higher cost savings.

4.5 Summary

A dynamic inventory policy can include stock reallocation to balance the inventory levels among the bases and as a result reduce the backorder cost. In this chapter stock reallocation policy is introduced as a general method for movement of parts among the bases, which covers both lateral transshipment and inventory rationing policies presented in the literature. The optimal stock reallocation problem is formulated as a dynamic program with the relative value function h^*. An approximation method using the state-space collapse phenomena is proposed. Since using the stock allocation policy results in the state-space collapse, the multi-dimensional queue length process can be approximated using a single queue system. Hence, in the proposed approximation, the relative value function of a K base model is approximated by the sum of the relative value functions of the aggregate queue and each individual single queue. In other words, the K base inventory system is treated as $K + 1$ independent queues with a queue representing the aggregate behavior of the inventory system and K queues representing the behavior of each individual base. Based on the numerical results, stock reallocation results in further cost savings. Additionally, the proposed approximation method is reliable in approximating the optimal stock reallocation policy.

Chapter 5

Emergency Resupply of Service Parts

5.1 Introduction

In the previous chapter, it was shown that further cost savings can be achieved by using more dynamic inventory policies. Stock reallocation policy allows parts transshipment among the bases and thus leads to decreased risk of backorder. Additionally, it determines when a stock reallocation action should be triggered, and the part should be sent from which base to the base in urgent need of service part. Assume all the initial stocks are at the repair facility being repaired and there are some backlogs at the bases. In this case, it is reasonable to find a part from other sources in order to fulfill the backlogged demands or make the repair facility flexible by allocating emergency repair servers. In the latter, emergency repair facility is activated with a one time setup cost and performs the repair job within negligible time. For the emergency resupply from other sources, it is assumed that a co-operation agreement has been signed between the MRO service provider and other MRO service provider(s). Such an agreement makes it possible to exchange a failed part with a ready-to-use one by paying a fixed amount of money.

When using the emergency resupply sources, a question arises: given a fixed emergency resupply cost, when is it optimal to activate the emergency repair server or request a ready-to-use part from other MRO service provider? This chapter answers this question by developing an emergency resupply policy. Such a policy, defines boundary lines in the region $\mathcal{R}_b$ of the state-space. When current state of the inventory system reaches the boundaries, it triggers an emergency resupply action and thus prevents the current state to pass it. Limiting boundaries are defined based on the optimal relative value function in the region $\mathcal{R}_b$. In this chapter, the exact value for h^* when $\mathbf{S} = 0$ is derived. Afterwards, based on the state-space collapse phenomenon mentioned in the previous chapter, approximate h^* under the emergency resupply policy is derived. Finally, computational experiments are performed to evaluate the efficiency of the proposed policy.

The remainder of this chapter is organized as follows: The problem under study is described and model formulation is presented in Section 5.2. In Section 5.3, the problem of optimal emergency resupply is addressed and the approximation method is described. Numerical experiments were conducted and the results were discussed in Section 5.4. Section 5.5 concludes with a summary of the contributions and findings of this chapter.

5.2 Problem Description and Model Formulation

The emergency resupply model developed in this chapter is built upon the model developed in the Chapter 3. However, some modifications are needed in order to address the emergency resupply policy. In this section, modifications made on the inventory system configurations and assumptions are explained and the problem formulation is presented.

5.2.1 System Description

Consider the repairable service parts inventory system described in Section 3.2. At the beginning of the inventory operations, initial stocks are provisioned and allocated to each base. During the operation, when a part is repaired, repair facility allocates the ready-to-use part to one of the bases based on the approximate stock allocation rule proposed in Section 3.3.2. When there is no inventory on-hand in the bases and there are some backorders in bases, it is reasonable to provide a service part from another source in the form of emergency resupply to clear the backorder(s), considering the fact that in commercial aircraft MRO, backorders are costly and uptime is crucial for customer satisfaction. The objective is to find the optimal time an emergency resupply is required and the base which receives the emergency resupply part. Figure 5.1 illustrates a graphical representation of the inventory network and emergency resupply of service parts within the resupply network.

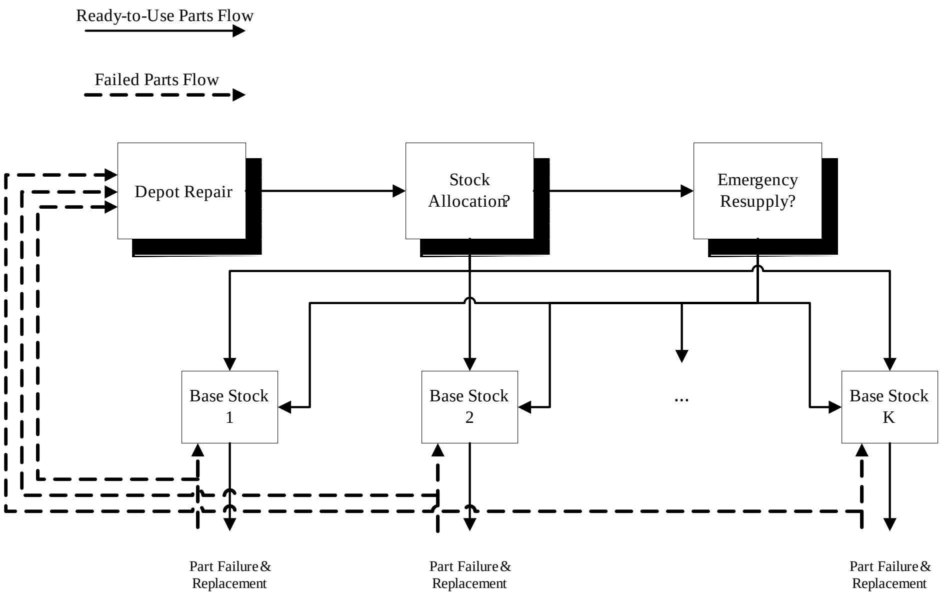


Figure 5.1: Service parts inventory system with emergency resupply

5.2.2 Notations and Assumptions

All of assumptions made in Section 3.2 applies to the model developed in this chapter. Additionally, it is assumed that emergency replenishment time is negligible and a one-time cost of π' per part is incurred in the case of emergency supply. Additionally, it is assumed that the borrowed service part is returned to the emergency resupply source, when all of the parts have been repaired and allocated to the bases. Hence, the base-stock levels at the bases remains constant.

5.2.3 MDP Formulation

Let $\mathbf{E}(t)$ denote the emergency resupply vector at time t, where $\mathbf{E}_k(t) = 1$ implies that there is an emergency resupply in base k and $\mathbf{E}_k(t) = 0$ otherwise.

The repair and allocation process at time step t can be defined by the following recursion:

$$\mathbf{X}(t) = \mathbf{X}(t-1) - B(t)\mathbf{F}(t-1) + \mathbf{A}(t) - \mathbf{A}(t) \times \mathbf{E}(t) \tag{5.1}$$

where the i.i.d process $(\mathbf{A}_k(1 \le k \le K), B)$ has marginal distribution defined by

$$\mathsf{P}\{(\mathbf{A}_1(t), \mathbf{A}_2(t), ..., \mathbf{A}_K(t), B(t)) = \mathbf{e}_i\} = \mathbf{p}_i, \tag{5.2}$$

with $\mathbf{p}_i = \Lambda_i$ for $i = 1, 2, ..., K$ and $\mathbf{p}_{K+1} = M$.

In this case, $(\mathbf{X}, \mathbf{E})$ is a Markov Decision Process, where the emergency resupply decision $\mathbf{E}(t)$ is defined based on the number of parts in the repair facility from each base. Let $\mathbf{G}^e(\mathbf{x})$ denote the emergency vector when the inventory state is $\mathbf{x}$. Emergency resupply cost under a particular emergency resupply e is:

$$\pi_e(\mathbf{x}) = \pi' \sum_{k=1}^{K} \Lambda_k \mathbf{G}^e(\mathbf{x}_k) \tag{5.3}$$

Let $\mathbf{c}_k$ denote the backorder cost at base k per unit per time. Assuming the current state is $\mathbf{x} = \mathbf{X}(t)$, the stage cost at time t is:

$$C_e(\mathbf{x}) = \pi_e(\mathbf{x}) + \sum_{k=1}^{K} \mathbf{c}_k [\mathbf{x}_k - \mathbf{S}_k]^+. \tag{5.4}$$

5.3 Emergency Resupply Policy

Let $h^*(\mathbf{x})$ and $\eta^*(\mathbf{S})$ denote the relative value function for the dynamic program and the steady-state average cost of the system under the optimal emergency resupply policy, respectively. The average cost is independent of the initial state

$\mathbf{x}$ and the dynamic programming optimality equation [135] can be defined as:

$$\eta^*(\mathbf{S})+h^*(\mathbf{x}) = \min_{e\in\mathsf{E}(\mathbf{x})}[C_e(\mathbf{x})+(\sum_{k=1}^{K}\Lambda_k h^*(\mathbf{x}+\mathbf{e}_k-\mathbf{G}^e(\mathbf{x}_k)))+Mh^*(\mathbf{x}-\mathbf{e}_f)]. \quad (5.5)$$

Let $\Phi"(x)$ denote the emergency resupply policy when current state is x. Optimal emergency resupply policy can be defined based on the optimal relative value function as follows:

$$\Phi''^*(\mathbf{x}) = \arg\min_{e\in\mathsf{E}(\mathbf{x})}[C_e(\mathbf{x}) + (\sum_{k=1}^{K}\Lambda_k h^*(\mathbf{x}+\mathbf{e}_k-\mathbf{G}^e(\mathbf{x}_k))) + Mh^*(\mathbf{x}-\mathbf{e}_f)]. \quad (5.6)$$

The optimal emergency supply policy is defined by boundaries in region $\mathcal{R}_b$ of the state-space. Figure 5.2 demonstrates such limiting boundary lines for the example described in Figure 3.3 with the emergency resupply cost of $\pi' = 50$. Consider the case, where the inventory level of this inventory system is $[-1,-1]$ or the inventory state is $\mathbf{x} = [17, 9]$ (i.e. each base has a backorder to fulfill). If the next event is a part failure at base 2, its inventory level drops to -2 and the inventory state becomes $\mathbf{x} = [17, 10]$. Since this inventory state is on the emergency resupply boundary line for base 2 or in other words $\mathbf{G}^e([17, 10]) = [01]$, a ready-to-use part is requested from the emergency resupply source. The part becomes available instantly and is sent to base 2 to fulfill one of its backorders and hence the inventory level changes to $[-1,-1]$.

In order to approximate the relative value function, the state-space collapse phenomena is used. The exact relative value function when $\mathbf{S} = 0$ is derived. Afterwards, a difference operator is defined based on the obtained relative value function. Finally, using the aggregate queue explained in the previous chapter

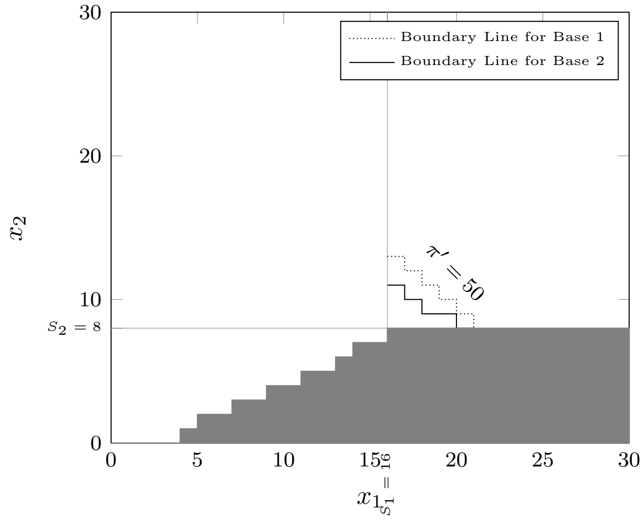


Figure 5.2: Optimal emergency resupply policy and boundary lines

together with the difference operator, approximate relative value function for the emergency resupply policy is derived.

5.3.1 Approximating h^*

To approximate the optimal boundaries, h^* in region $\mathcal{R}_b$ needs to be approximated. Consider an inventory system with zero initial stock. That is an inventory system with region $\mathcal{R}_b$ as the only region in the state-space. For an inventory system with zero initial stocks, the relative value function h_0^* can be expressed as:

$$h_0^*(\mathbf{x}) = 1/2(\mathbf{x}\mathbf{D}\mathbf{x}^T + \mathbf{D}'\mathbf{x}), \tag{5.7}$$

where $\mathbf{D}$ is defined as follows:

$$\mathbf{D}_{i,j} = \mathbf{D}_{k,k}, \qquad k = \min\{i, j\},$$

$$\mathbf{D}_{i,i} = \frac{\mathbf{c}_i + \sum_{k=1}^{i-1} \mathbf{D}_{k,k}\lambda_k}{\mu - \sum_{k=i}^{K} \lambda_k}, \tag{5.8}$$

and $\mathbf{D}'$ is a vector where $\mathbf{D}'_i = \mathbf{D}_{i,i}$.

To prove that the relative value function can be expressed as proposed, one should prove that $\mathcal{D}h_0^*(\mathbf{x}) = -C(\mathbf{x}) + \eta^*$ holds for the proposed relative value function. Rescale the time so $\sum_{k=1}^{K} \lambda_k + \mu = 1$. $\mathcal{D}h_0^*(\mathbf{x})$ can be expressed as follows:

$$\begin{aligned} \mathcal{D}h_0^*(\mathbf{x}) &= P_f h_0^*(\mathbf{x}) - (\lambda + \mu)h^*(\mathbf{x}) \\ &= \mu(h_0^*(\mathbf{x} - \mathbf{e}_j) - h_0^*(\mathbf{x})) + \sum_{k=1}^{K} \lambda_k (h_0^*(x + e_k) - h_0^*(x)). \end{aligned} \tag{5.9}$$

Based on Equation 5.7 the following equation holds:

$$h_0^*(\mathbf{x} - \mathbf{e}_j) - h_0^*(\mathbf{x}) = -\sum_{k=1}^{K} \mathbf{D}_{k,k}\mathbf{x}_k \tag{5.10}$$

$$h_0^*(\mathbf{x} + \mathbf{e}_k) - h_0^*(\mathbf{x}) = \sum_{i=1}^{k-1} \mathbf{D}_{i,i}\mathbf{x}_i + \mathbf{D}_{k,k}(\sum_{j=k}^{K} \mathbf{x}_j) + \mathbf{D}_{k,k}. \tag{5.11}$$

Hence, $\mathcal{D}h_0^*(x)$ can be expressed as:

$$\mathcal{D}h^*(\mathbf{x}) = \mu(-\sum_{k=1}^{K} \mathbf{D}_{k,k}\mathbf{x}_k) + \sum_{k=1}^{K} \lambda_k (\sum_{i=1}^{k-1} \mathbf{D}_{i,i}\mathbf{x}_i + \mathbf{D}_{k,k}(\sum_{j=k}^{K} \mathbf{x}_j) + \mathbf{D}_{k,k}) \tag{5.12}$$

Taking out the factor $\mathbf{D}_{k,k}$ as the common factor:

$$\mathcal{D}h_0^*(\mathbf{x}) = \lambda_k (\sum_{k=1}^{K} \mathbf{D}_{k,k}\mathbf{x}_k (-\mu + \sum_{i=k}^{K}) + \sum_{k=1}^{K} \mathbf{D}_{k,k}) \tag{5.13}$$

For each k:

$$\mathbf{D}_{k,k}\mathbf{x}_k(-\mu+\sum_{i=k}^{K})\lambda_k = \frac{\mathbf{c}_i+\sum_{i=1}^{k-1}\mathbf{D}_{i,i}\lambda_i}{\mu-\sum_{i=k}^{K}\lambda_k}\mathbf{x}_k(-\mu+\sum_{i=k}^{K}) = -\mathbf{c}_k \tag{5.14}$$

From other side optimal average cost can be expressed as:

$$\eta^* = \sum_{k=1}^{K}\mathbf{D}_{k,k}\lambda_k \tag{5.15}$$

Putting Equations 5.14 and 5.15 into Equation 5.13 results in:

$$\begin{aligned}\mathcal{D}h_0^*(x) &= \eta^* - \sum_{k=1}^{K}\mathbf{c}_k \\ &= \eta^* + C(\mathbf{x})\end{aligned} \tag{5.16}$$

Let i and j denote two bases where $j = i+1$. Let $\mathcal{F}_j$ denote an operator defined on h as follows:

$$\mathcal{F}_j h(\mathbf{x}) = h(\mathbf{x}) - h(\mathbf{x}+\mathbf{e}_i-\mathbf{e}_j) \tag{5.17}$$

Consider two inventory systems, one with and the other without initial stocks. Let $\mathbf{x}$ denote a state in the inventory system with initial stocks where $\mathbf{x} \succeq \mathbf{S}$ and $\mathbf{y}$ denote a state in the inventory system without initial stocks, where $\mathbf{y} = \mathbf{x} - \mathbf{S}$. Based on the property of optimal stock allocation policy presented in Section 3.3.1, the optimal stock allocation policy for the two inventory systems while there is no stock on-hand is the same. Hence, assuming the stock allocation critical level for the lower base is zero,the following relationship holds for the relative value function of the two inventory systems:

$$\mathcal{F}_j h^*(\mathbf{x}) = \mathcal{F}_j h_0^*(\mathbf{y}) \tag{5.18}$$

Based on this relationship, $h^*(\mathbf{x})$ for the inventory system with initial stocks can be derived, provided that $h^*(\mathbf{x})$ where $\mathbf{x} - \mathbf{S} = n\mathbf{e}_k$ and $n \geq 0$ for some base k is available. However, finding $h^*(x)$ for such a condition is intractable. It is proposed to approximate the $h^*(\mathbf{x})$ where $\mathbf{x} - \mathbf{S} = n\mathbf{e}_1$. As mentioned in Chapter 4, stock allocation process results in the state-space collapse and the queue length process can be approximated by treating the stock allocation switching curve as the aggregate queue. Based on the state-space collapse, one can approximate the relative value function when the inventory level at all of the bases are zero except the base with lowest backorder cost i.e.($\mathbf{x}_k = \mathbf{S}_k, k = 2...K$). The relative value function can be approximated using the aggregate relative value function $h_b^*(\bar{x})$:

$$\tilde{h}(\mathbf{x}) = h_b^*(\bar{\mathbf{x}}), \qquad \mathbf{x} - \mathbf{S} = n\mathbf{e}_1, \quad n \geq 0 \tag{5.19}$$

However, when there is emergency resupply, aggregate queue is capacitated. In order to find the average cost of the aggregate queue, its optimal size needs to be found. Let l'^* denote the optimal size of the aggregate queue. l'^* minimizes the following average cost:

$$\eta = \frac{\pi' \rho^{l'^*}(1-\rho)\sum_{k=1}^{K}\Lambda_k}{1-\rho^{l'^*+1}} + \frac{\rho^{l'^*}(l'^*\rho - l'^* - 1) + \rho}{(1-\rho)(1-\rho^{l'^*+1})}. \tag{5.20}$$

Let $\tilde{h}(\mathbf{x})$ denote the approximate relative value function in region $\mathcal{R}_b$ for an inventory system with initial stocks and $\mathcal{F}_j \tilde{h}(\mathbf{x}) = \mathcal{F}_j h_0^*(\mathbf{x})$. In this case, $\tilde{h}(\mathbf{x})$

can be defined as follows:

$$\tilde{h}(\mathbf{x}) = h_0^*(\mathbf{y}) - h_0^*(\bar{\mathbf{y}}\mathbf{e}_1) + \hat{h}^*(\mathbf{x}), \qquad \mathbf{x} >= \mathbf{S}, \quad \mathbf{y} = \mathbf{x} - \mathbf{S}. \tag{5.21}$$

After approximating the $h^*(x)$ under the emergency resupply policy, approximate emergency resupply policy can be defined easily as follows:

$$\tilde{\Phi}''(\mathbf{x}) = \arg\min_{e \in \mathsf{E}(\mathbf{x})} [C_e(\mathbf{x}) + (\sum_{k=1}^{K} \Lambda_k \tilde{h}(\mathbf{x} + \mathbf{e}_k - \mathbf{G}^e(\mathbf{x}_k))) + M\tilde{h}(\mathbf{x} - \mathbf{e}_f)]. \tag{5.22}$$

Proposed policy approximates the future possible backorders using the relative value function. In case of a failure in one of the bases, if the increase in the backorder risk exceeds the emergency resupply cost, it initiates an emergency resupply to that base. As can be seen, emergency resupply policy is defined based on the vector e, which is the minimizer of the expression presented in Equation 5.22 and has the same properties as $\mathbf{E}(t)$ presented in Section 5.2.3. If $e = 0$, it implies that no emergency resupply of service part is needed. However, if $e_k = 1$, an emergency resupply of service part is initiated by base k.

5.4 Numerical Study: Results and Discussion

In order to evaluate the accuracy of the proposed approximate emergency resupply boundaries and to investigate the amount of cost savings achieved by emergency resupply policy, computational experiments are conducted. Average cost is used as the performance measure and is calculated using the value iteration method. The optimality gaps of the proposed emergency resupply policy is investigated and in order to obtain further insights on the performance of the

Table 5.1: Input parameters for general evaluation of emergency resupply policies

Parameters	**Values**
λ_2/λ_1	1/3,1/2,1,2,3
c_2/c_1	1,2,3
ρ	0.8
Total Stock	8,12
π'	30,50

approximate policy, sensitivity analysis is carried out.

To meet the first objective of the numerical experiment, a two-base model is considered and a test bed with a factorial design on five parameters is used. Stock allocation policy and initial base-stock levels is obtained using the method described in Section 3.3.2. For the second objective, two experiments are designed. To investigate the impact of different parameters except the number of bases on the optimality gap, a two-base model is considered and the performance of each policy is evaluated by varying each of the parameters. The test bed and average cost calculation method are similar to the previous experiment. Since real world service parts supply chain constitute higher number of bases, another experiment is carried out, in which the number of bases is varied from 3 to 10 and a simulation study is conducted to compare the performance of the proposed emergency resupply policy with the inventory system without emergency resupply capability.

To investigate the optimality gap of the approximate emergency resupply policy, a test bed is used based on factorial design. Parameter values are presented in Table 5.1.

Combining all the parameters 52 problem instances are defined. Numerical results are summarized in Table A.4. It was observed that introducing the emergency resupply capability to the inventory system resulted in 37.144% reduction in

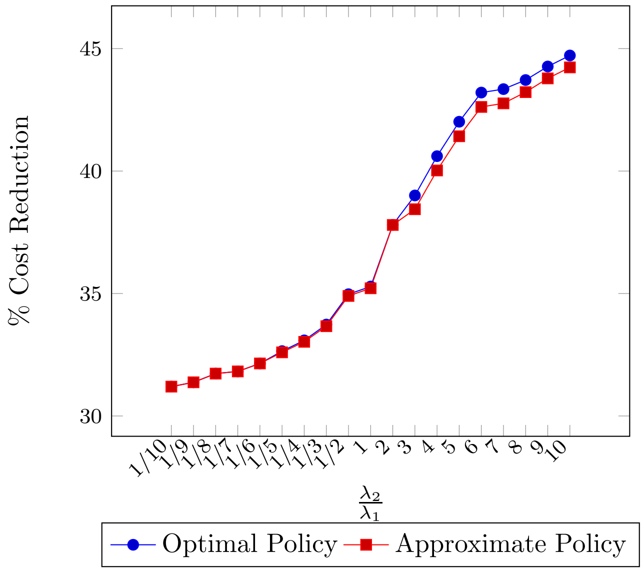


Figure 5.3: Demand sensitivity analysis

cost. Additionally, Approximate emergency resupply policy performed very close to the optimal policy with maximum and average optimality gaps of 1.732% and 0.343%, respectively.

Sensitivity Analysis In addition to the problem sets described before,sensitivity analysis is conducted to investigate the effect of the each input parameter on the optimality gap. Sensitivity analysis results are as follows:

- **Demand Ratio:** Demand rates are varied between 1 to 10, resulting in a demand ratio of 0.1 to 10. For each case backorder cost is varied as $c = \{1, 2, 3\}$, while ρ is 0.8, total initial stock is set to 10 and emergency resupply cost is 40. It was observed that as the demand ratio increases, the amount of cost savings and the optimality gaps increases (Figure 5.3). Since $c_2/c_1 \geq 1$, as the demand ratio increases, there are more instances of emergency resupply to fulfill the backorders of base 2.

- **Backorder Cost:** Backorder costs are set to $c_1 = 1, c_2 = \{1, 2, ..., 10\}$.

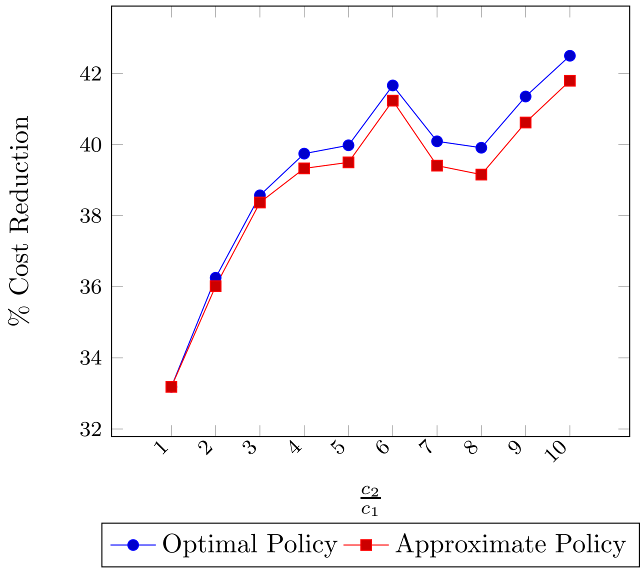


Figure 5.4: Backorder cost sensitivity analysis

Demand rates are varied between $\{1, 2, 3\}$ and other parameters are the same as the demand ratio sensitivity analysis. Based on the numerical results as c_2 increases, the amount of cost savings and the optimality gaps increases (Figure 5.4). Since the emergency resupply cost is 40 for all of the experiments, higher backorder costs at base 2 leads to more frequent emergency resupplies to this base.

- **Repair Server Utilization:** Repair server utilization is set to be $\rho = \{0.5, 0.6, 0.7, 0.8, 0.9, 0.95\}$, while $c_1 = 1, c_2 = \{1, 2, 3\}$ and all other parameters are the same as the previous experiment. It was observed that as server utilization increase, cost reduction from using the emergency resupply source increase, too. As the server utilization increase, inventory state is highly concentrated in region $\mathcal{R}_b$ and hence more cost savings can be achieved by implementation of the emergency resupply policy. Additionally, no particular trend was observed in terms of the optimality gaps when

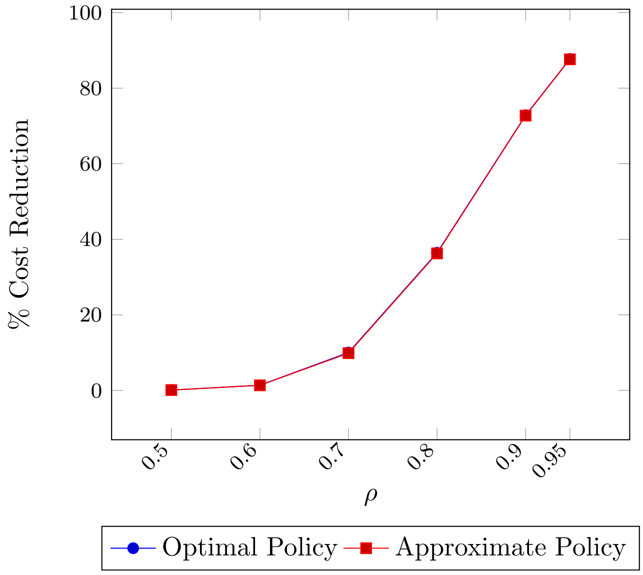


Figure 5.5: Repair server utilization sensitivity analysis

the server utilization changed (Figure 5.5).

- **Investment Budget:** Initial stock is varied between $\{6, 10, 14, 18, 22, 26\}$, while repair server utilization is set to be 0.8 and all other parameters are the same as the previous experiment. It was observed that as the base-stock level increase, the amount of cost reduction increase (See Figure 5.6).

- **Emergency Resupply Cost:** Emergency resupply cost is varied between $\{20, 30, 40, 50, 60, 70\}$, while initial total stock is set to be 10 and all other parameters are the same as the previous experiment. It was observed that as the emergency resupply cost increases, cost reduction achieved by emergency resupply decreases exponentially. As the emergency resupply cost increases, the frequency of emergency resupply decreases. The accuracy of the approximate policy is insensitive to the emergency resupply cost (See

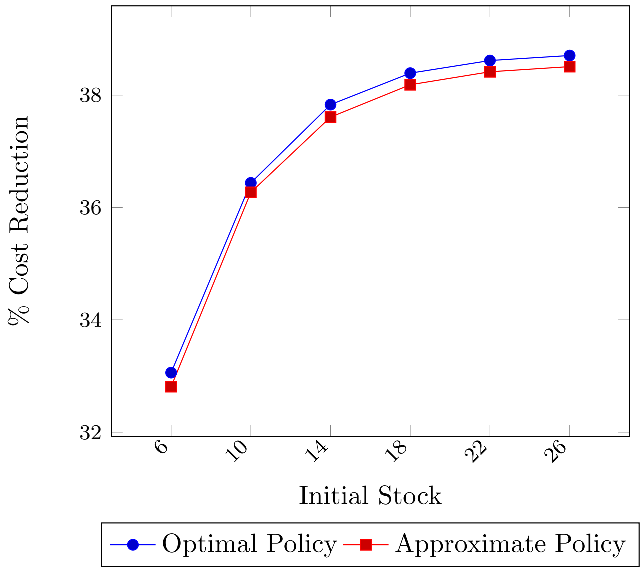


Figure 5.6: Initial budget sensitivity analysis

Figure 5.7).

- **Number of Bases:** To investigate the amount of cost reductions from using the proposed emergency resupply policy for large number of bases, a simulation study is conducted. Inventory systems of 3 to 10 bases are considered. For each number of base, the initial base-stock level is set to have an expected cost of around to 0.1 and 0.01 and two values of 30 and 50 are considered as the emergency resupply cost. For each case, The combination of initial base-stock level and emergency resupply policy results in 4 combinations. For each combination, 100 problem sets are created with random demand rates and backorder costs. Demand rates are drawn from uniform distribution between $(0, K]$. It is assumed that there are three classes of customers with backorder costs of $\{1, 2, 3\}$ and each base is assigned to one of the classes, randomly. For each problem set, two models of inventory with emergency resupply policy and without emergency

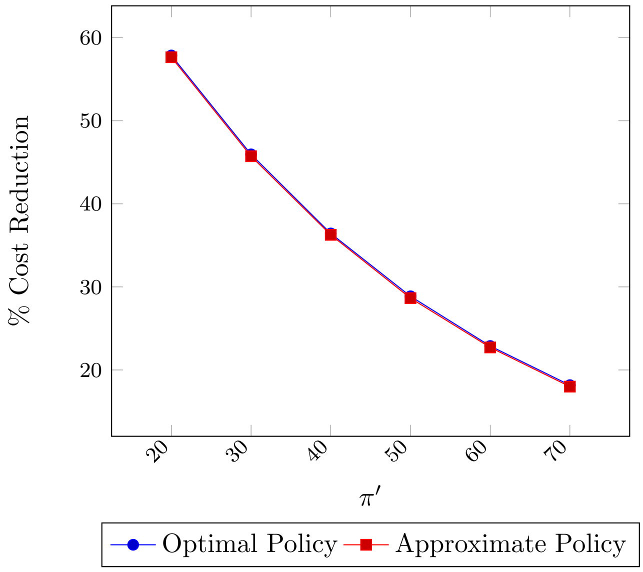
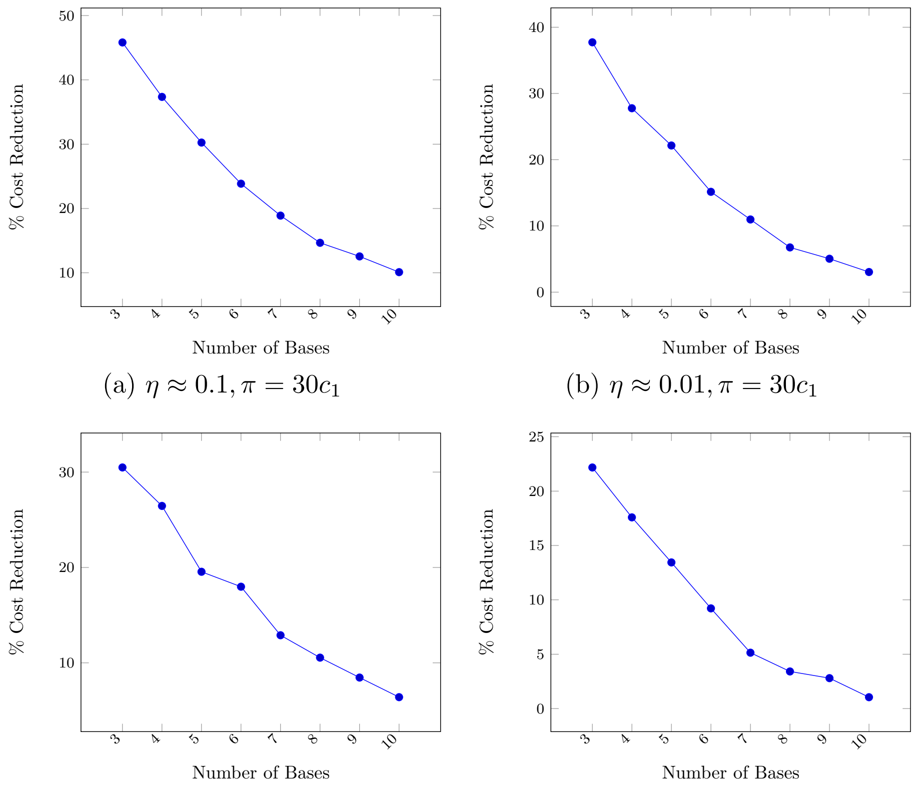


Figure 5.7: Emergency resupply cost sensitivity analysis

resupply policy are simulated for 5 replications of 500000 time steps. A warm-up period of 500 time steps is also considered. Figure 5.8 illustrate the simulation result. Relative difference in each case is calculated by average backorder cost of 100 random problem sets for the models with and without emergency resupply policy.

As can be seen, as the number of bases increase, cost reduction from using the approximate emergency resupply policy decreases. Consider the case where the current state ($\mathbf{x}$) is inside the region $\mathcal{R}_b$ but no emergency resupply has been initiated and if a part fails at base k, it is fulfilled by an emergency resupply. The next event is a demand at base k with probability Λ_k. As the number of bases increase, Λ_k decreases and there is a higher probability that there will be still backorder(s) at the bases without any emergency resupply.



(a) $\eta \approx 0.1, \pi = 30c_1$

(b) $\eta \approx 0.01, \pi = 30c_1$

(c) $\eta \approx 0.1, \pi = 50c_1$

(d) $\eta \approx 0.01, \pi = 50c_1$

Figure 5.8: Number of bases sensitivity analysis

5.5 Summary

When there are emergency resupply sources, MRO service providers can reduce the backorder cost by fulfilling the demand from those sources. Such sources can be in the form of emergency repair servers or other MRO service providers with co-operation agreements. In this chapter, an inventory policy is developed to find out when is the optimal time to fulfill the demand using emergency resupply sources. The optimal emergency resupply policy is modeled as a dynamic program with the relative value function h^*. The exact value of h^* when the initial stocks are zero is derived. Afterwards, using this value together with the state-space collapse phenomenon, the relative value function in the inventory system with initial stocks is approximated. Based on the numerical results, it is observed that the amount of cost savings achieved by emergency resupply policy is much higher than those for stock allocation and reallocation policies. Additionally, it is observed that the approximate policy performs very close to the optimal policy.

Chapter 6

Conclusion

6.1 Research Summary

Repairable service parts supply chain has been researched since 1960s. Researchers in this broad field have focused on different aspects of the service parts supply chain, such as supply process, demand process, supplier/customer relationships and the structure of service parts and its effects in the supply chain management. Inventory control models form a large part of the literature in this field. Earlier models were developed for military systems, and their main objective was to find the steady-state probability distribution of the inventory level for a particular inventory policy. Main assumptions of these models were the $(S-1, S)$ ordering policy with Firs-In-First-Out (FIFO) dispatching policy in a two-echelon resupply network, which was suitable for the military system. The steady-state performance measure was used to optimize the number of initial stocks subject to a performance measure constraint. Subsequent research was done trying to relax some of the main assumptions of the earlier models and hence improving the performance measure of the inventory system. While earlier models had assumed the order for part replenishment can only be replenished by the higher echelon, a stream of research focused on adding neighboring bases as an alternative source for ordering the service parts. Another stream tried to relax the FIFO dispatching policy and make real-time decisions on parts allocation to the bases. These models were able to improve the performance of the inventory system by introducing more flexible inventory policies.

One-for-one replenishment policy and exponential failure and repair times have made Markov chain and queueing theory a dominant model development and analysis tool in most of the literature in service parts inventory models. Research issues in applying the queueing theory such as the so-called curse-of-dimensionality issue have limited the developed models in the literature in terms of the flexibility of the policies and the size of network they can be applied to. However, recent changes in commercial aviation have led to more decentralized and complex service parts supply chain as many airlines have outsourced their MRO operations to other airlines or an independent third party MRO service provider. This research investigates dynamic inventory policies for large service parts re-supply networks which capitalizes on real-time data related to the number of parts at each location within the resupply network.

Three dynamic inventory policies are developed, namely stock allocation, stock reallocation and emergency resupply. Stock allocation policy involves the dynamic dispatching decision made by the repair facility. It relaxes the FIFO dispatching rule considered by most of the literature. With stock allocation policy, the repair facility is able to identify the base with the most urgent need for part and send the repaired part accordingly. Stock reallocation policy involves the movement of parts among the bases in order to decrease the backorder risk. Movement of parts can be physical or virtual, and hence it represents the lateral transshipment and inventory rationing policies developed in the literature. With stock reallocation policy, the decision maker is able to identify when a reallocation is needed, and which base should send a part to which base. In addition to the service parts within the resupply network, in case of stock-out in the resupply network, one might need to secure a part from other sources in order to decrease the number of backorders. Emergency resupply policy makes it possible for the decision maker to find out when a part should be asked from other sources and where it should be sent. To develop these policies, this research has tried to address the common problem with Markov Decision Process (MDP), i.e. curse of dimensionality, by developing heuristic approximation methods for the respective optimal relative value functions. Approximation methods are mainly based on the relative value function of a single inventory model.

As a starting point in dynamic policy generation, the stock allocation problem of repaired parts to the bases is addressed. Related dynamic program is modeled as a MDP. The structure of the optimal policy is investigated by finding the optimal policy for a two-base model using the value iteration method. It is shown that the stock allocation switching curve passes the optimal initial base-stock level and the solution to the optimal initial base-stock level can be found by finding the optimal stock allocation policy and vice versa. A stock allocation method is proposed based on the combination of the value function and average backorder cost of a single base. For an inventory system with two bases, it is completely based on the relative value function of a single base and as the number of bases increase, it relies more on the average backorder cost of a single base. Two of these regions are of high importance: when all the bases have stock ($\mathcal{R}_s$) and when all of them have backorders ($\mathcal{R}_b$). Based on the numerical results, approximate stock allocation policy performs close to the optimal policy.

In order to develop stock reallocation and emergency resupply policies, approximate relative value functions in regions $\mathcal{R}_s$ and $\mathcal{R}_b$ are needed. Stock allocation policy leads to a phenomenon called "state-space collapse" in the literature. In this case, the multi-dimensional state-space queue length can be approximated using a single aggregate queue length with high accuracy. This fact is used in Chapter 4 to develop aggregate queues based on the stock allocation policy. A single inventory model is considered and the optimal relative value function is derived by solving the respective difference equation systems. Afterwards, approximate relative value function in region $\mathcal{R}_s$ is derived based on the relative value function of a single base. In the proposed approximation method, the relative value function of a K base model is approximated by the sum of the relative value functions of the aggregate queue and each individual single queue. Based on the numerical results, the proposed approximation method is reliable in approximating the optimal hedging lines, which define the stock reallocation policy.

To approximate the relative value function in region $\mathcal{R}_b$, it is first assumed that none of the bases hold initial stocks. By solving the related difference equation systems, relative value function of such an inventory model is derived. From the other side, the aggregate queue resulted from the state-space collapse in region $\mathcal{R}_b$ is placed in the path where the base with lowest backorder cost has backorders while other bases has no on-hand inventory or backorders (This fact is based on the stock allocation policy in this region i.e. send the repaired part to the base with highest backorder cost). The optimal size of the aggregate queue on this path is obtained and its relative value function based on the optimal size is derived. Afterwards, the approximate relative value function on this path is set to be equal to the relative value function of the capacitated aggregate queue and the approximate relative value function in region $\mathcal{R}_b$ is derived based on the value difference of the system with zero stocks and the approximate relative value function on the stock allocation path in this region.

For each of the developed inventory policies, computational experiments were conducted. Computational experiments of each policy consisted a general problem set and series of sensitivity analysis. Based on the numerical results, proposed approximate policies performed close to the optimal policies.

6.2 Research Contributions

In this research three questions mentioned in Section 1.3 have been answered. Contributions of this research to the literature of service parts supply chain can be summarized as follows:

- Based on queueing theory and Markov Decision Process, the relative value function of a single base model is derived.
- An approximate stock allocation policy is developed based on the relative value function and average cost of a single base model. The proposed policy performed very close to the optimal policy and outperformed the myopic policy an all of the cases.
- A stock reallocation policy is developed based on the approximate relative value function in the region $\mathcal{R}_s$.

- An emergency resupply policy is developed as an additional option of supply. It has shown that an optimal resupply policy results in limiting boundaries in region $\mathcal{R}_b$. Approximate limiting boundaries have been developed by approximating the optimal relative value function in this region.

Approximate relative value functions can be used to generate dynamic inventory policies, where the real-time information of service parts within the resupply network is available. Proposed techniques are able to handle decision making in the resupply networks with large number of bases, where the traditional optimization techniques such as the value-iteration method are computationally intensive. Moreover, dynamic inventory policies presented in this research result in look-up tables, which can be implemented in decision support systems easily.

6.3 Future Work

Following future directions of research are proposed:

1. In this research, single-echelon inventory model is considered. However, proposed approaches can be utilized in multi-echelon networks. This can be done by deriving the expected waiting times to replenish an order of the bases of the lower echelon by the base in the higher echelon. Afterwards, relative value function for a single base model with infinite repair servers can be derived, the repair time being the average waiting time. Relative value function for the single location model can be the basis for approximating the relative value functions of the models with higher number of bases.

2. Simplistic assumptions were made in the development of emergency resup-

ply policy. More research is needed to develop co-operation mechanism between MRO service providers. One approach could be to introduce the new concept of service parts renting, where instead of exchanging the failed part with a ready-to-use one, a MRO service provider rents the service part and is charged for the length of time it has borrowed the part. Another approach is to define time windows for the return of the part. Such a problem is more complex as one needs to find the optimal time to return the part in addition to the optimal time for initiating the emergency resupply. Developing pricing techniques and co-operative strategies for these models is another possible direction for research.

3. One-for-one replenishment is common for the slow-moving items such as MRO service parts. Proposed approach can be generalized to develop dynamic inventory policies under the batch ordering policy. Such a model can be addressed by a queueing network with switching costs, i.e. when the base to receive the repaired part is changed, an additional cost is charged due to the set-up cost for order fulfillment.

4. Although only unscheduled part failures were considered, scheduled part replacements are also done by MRO service providers. Combining the proposed model with scheduled demand for parts is another interesting research.

Appendix A

Tables

Table A.1: Optimality gaps of stock allocation policies

ρ	Total Stock	λ	Backorder Cost	η^*	Optimality Gap (%)	
					VF-Based Policy	Myopic Policy
0.8	8	(1,1)	(1,1)	0.702	0.000	0.000
0.8	8	(1,2)	(1,1)	0.700	0.000	0.309
0.8	8	(1,3)	(1,1)	0.700	0.078	1.330
0.8	8	(3,1)	(1,2)	0.754	0.492	0.896
0.8	8	(2,1)	(1,2)	0.770	0.000	0.130
0.8	8	(1,1)	(1,2)	0.819	0.000	0.000
0.8	8	(1,2)	(1,2)	0.892	0.000	0.618
0.8	8	(1,3)	(1,2)	0.940	0.000	1.036
0.8	8	(3,1)	(1,3)	0.798	0.332	0.220
0.8	8	(2,1)	(1,3)	0.832	0.123	0.470
0.8	8	(1,1)	(1,3)	0.907	0.000	0.859
0.8	8	(1,2)	(1,3)	1.023	0.000	2.286
0.8	8	(1,3)	(1,3)	1.126	0.080	2.883
0.8	12	(1,1)	(1,1)	0.289	0.000	0.000
0.8	12	(1,2)	(1,1)	0.288	0.095	0.746
0.8	12	(1,3)	(1,1)	0.288	0.165	1.242
0.8	12	(3,1)	(1,2)	0.310	0.657	1.098
0.8	12	(2,1)	(1,2)	0.317	0.000	0.291
0.8	12	(1,1)	(1,2)	0.337	0.000	0.000
0.8	12	(1,2)	(1,2)	0.367	0.398	0.636
0.8	12	(1,3)	(1,2)	0.388	0.586	2.271
0.8	12	(3,1)	(1,3)	0.329	0.378	0.429
0.8	12	(2,1)	(1,3)	0.342	0.375	0.205
0.8	12	(1,1)	(1,3)	0.373	0.000	0.826
0.8	12	(1,2)	(1,3)	0.422	0.039	2.087
0.8	12	(1,3)	(1,3)	0.464	0.246	4.911
0.9	8	(1,1)	(1,1)	3.907	0.015	0.015
0.9	8	(1,2)	(1,1)	3.904	0.187	0.107
0.9	8	(1,3)	(1,1)	3.904	0.414	0.317
0.9	8	(3,1)	(1,2)	4.041	0.121	0.215
0.9	8	(2,1)	(1,2)	4.097	0.120	0.038
0.9	8	(1,1)	(1,2)	4.262	0.014	0.014
0.9	8	(1,2)	(1,2)	4.484	0.020	0.238
0.9	8	(1,3)	(1,2)	4.689	0.048	0.466
0.9	8	(3,1)	(1,3)	4.167	0.140	0.059
0.9	8	(2,1)	(1,3)	4.278	0.062	0.086
0.9	8	(1,1)	(1,3)	4.478	0.013	0.370
0.9	8	(1,2)	(1,3)	4.883	0.062	0.590
0.9	8	(1,3)	(1,3)	5.225	0.125	0.893
0.9	12	(1,1)	(1,1)	2.564	0.028	0.028
0.9	12	(1,2)	(1,1)	2.562	0.234	0.197
0.9	12	(1,3)	(1,1)	2.562	0.398	0.335
0.9	12	(3,1)	(1,2)	2.652	0.231	0.264
0.9	12	(2,1)	(1,2)	2.689	0.146	0.085
0.9	12	(1,1)	(1,2)	2.797	0.027	0.027
0.9	12	(1,2)	(1,2)	2.944	0.084	0.239
0.9	12	(1,3)	(1,2)	3.079	0.109	0.738
0.9	12	(3,1)	(1,3)	2.735	0.252	0.109
0.9	12	(2,1)	(1,3)	2.808	0.136	0.061
0.9	12	(1,1)	(1,3)	2.940	0.025	0.368
0.9	12	(1,2)	(1,3)	3.207	0.085	0.511
0.9	12	(1,3)	(1,3)	3.433	0.194	1.328

Table A.2: Performance comparison of stock reallocation policies ($\tau = 5\mu$)

π	Total Stock	λ	Backorder Cost	η^*	Cost Reduction (%)	
					Optimal Policy	Approximate Policy
0.01	8	(1,1)	(1,1)	0.674	4.008	4.008
		(1,2)	(1,1)	0.673	3.858	3.858
		(1,3)	(1,1)	0.673	3.936	3.929
		(3,1)	(1,2)	0.717	5.382	5.315
		(2,1)	(1,2)	0.735	4.571	4.330
		(1,1)	(1,2)	0.785	4.224	4.218
		(1,2)	(1,2)	0.857	3.941	3.933
		(1,3)	(1,2)	0.904	3.914	3.905
		(3,1)	(1,3)	0.758	5.359	5.001
		(2,1)	(1,3)	0.795	4.520	4.511
		(1,1)	(1,3)	0.869	4.151	4.146
		(1,2)	(1,3)	0.984	3.784	3.770
		(1,3)	(1,3)	1.082	3.975	3.882
	12	(1,1)	(1,1)	0.276	4.403	4.403
		(1,2)	(1,1)	0.276	4.324	4.322
		(1,3)	(1,1)	0.276	4.522	4.436
		(3,1)	(1,2)	0.294	6.031	5.965
		(2,1)	(1,2)	0.301	5.141	4.895
		(1,1)	(1,2)	0.321	4.618	4.612
		(1,2)	(1,2)	0.351	4.777	4.771
		(1,3)	(1,2)	0.370	5.039	5.029
		(3,1)	(1,3)	0.310	5.873	5.505
		(2,1)	(1,3)	0.326	5.248	5.237
		(1,1)	(1,3)	0.356	4.588	4.582
		(1,2)	(1,3)	0.403	4.568	4.556
		(1,3)	(1,3)	0.443	4.720	4.612
0.05	8	(1,1)	(1,1)	0.675	3.959	3.959
		(1,2)	(1,1)	0.674	3.797	3.763
		(1,3)	(1,1)	0.674	3.839	3.247
		(3,1)	(1,2)	0.717	5.319	5.200
		(2,1)	(1,2)	0.735	4.516	4.242
		(1,1)	(1,2)	0.785	4.164	4.164
		(1,2)	(1,2)	0.857	3.863	3.857
		(1,3)	(1,2)	0.904	3.841	3.777
		(3,1)	(1,3)	0.758	5.297	4.922
		(2,1)	(1,3)	0.796	4.463	4.355
		(1,1)	(1,3)	0.870	4.067	4.067
		(1,2)	(1,3)	0.985	3.695	3.623
		(1,3)	(1,3)	1.083	3.880	3.723
	12	(1,1)	(1,1)	0.276	4.347	4.329
		(1,2)	(1,1)	0.276	4.254	4.220
		(1,3)	(1,1)	0.276	4.417	4.346
		(3,1)	(1,2)	0.294	5.959	5.841
		(2,1)	(1,2)	0.301	5.077	4.805
		(1,1)	(1,2)	0.322	4.553	4.550
		(1,2)	(1,2)	0.351	4.686	4.680
		(1,3)	(1,2)	0.371	4.941	4.848
		(3,1)	(1,3)	0.311	5.805	5.362
		(2,1)	(1,3)	0.326	5.180	5.073
		(1,1)	(1,3)	0.356	4.500	4.490
		(1,2)	(1,3)	0.404	4.468	4.396
		(1,3)	(1,3)	0.444	4.607	4.451

Table A.3: Performance comparison of stock reallocation policies ($\tau = 10\mu$)

π	Total Stock	λ	Backorder Cost	η^*	Cost Reduction (%)	
					Optimal Policy	Approximate Policy
0.01	8	(1,1)	(1,1)	0.673	4.213	4.213
		(1,2)	(1,1)	0.672	4.020	4.020
		(1,3)	(1,1)	0.672	4.098	4.044
		(3,1)	(1,2)	0.715	5.614	5.576
		(2,1)	(1,2)	0.734	4.757	4.599
		(1,1)	(1,2)	0.784	4.344	4.344
		(1,2)	(1,2)	0.856	4.065	4.065
		(1,3)	(1,2)	0.903	4.021	4.014
		(3,1)	(1,3)	0.756	5.538	5.307
		(2,1)	(1,3)	0.794	4.648	4.647
		(1,1)	(1,3)	0.868	4.286	4.286
		(1,2)	(1,3)	0.983	3.880	3.878
		(1,3)	(1,3)	1.081	4.077	4.012
	12	(1,1)	(1,1)	0.276	4.608	4.608
		(1,2)	(1,1)	0.275	4.490	4.490
		(1,3)	(1,1)	0.275	4.692	4.638
		(3,1)	(1,2)	0.293	6.266	6.228
		(2,1)	(1,2)	0.300	5.328	5.171
		(1,1)	(1,2)	0.321	4.740	4.739
		(1,2)	(1,2)	0.350	4.908	4.908
		(1,3)	(1,2)	0.370	5.168	5.107
		(3,1)	(1,3)	0.310	6.054	5.823
		(2,1)	(1,3)	0.325	5.379	5.375
		(1,1)	(1,3)	0.355	4.723	4.723
		(1,2)	(1,3)	0.403	4.669	4.668
		(1,3)	(1,3)	0.443	4.811	4.744
0.05	8	(1,1)	(1,1)	0.673	4.187	4.187
		(1,2)	(1,1)	0.673	3.984	3.960
		(1,3)	(1,1)	0.673	4.038	3.383
		(3,1)	(1,2)	0.715	5.581	5.462
		(2,1)	(1,2)	0.734	4.727	4.499
		(1,1)	(1,2)	0.784	4.313	4.313
		(1,2)	(1,2)	0.856	4.018	3.977
		(1,3)	(1,2)	0.903	3.982	3.914
		(3,1)	(1,3)	0.757	5.508	5.219
		(2,1)	(1,3)	0.794	4.619	4.537
		(1,1)	(1,3)	0.868	4.241	4.192
		(1,2)	(1,3)	0.984	3.828	3.780
		(1,3)	(1,3)	1.081	4.016	3.918
	12	(1,1)	(1,1)	0.276	4.580	4.580
		(1,2)	(1,1)	0.275	4.450	4.400
		(1,3)	(1,1)	0.276	4.627	4.558
		(3,1)	(1,2)	0.293	6.229	6.112
		(2,1)	(1,2)	0.301	5.295	5.068
		(1,1)	(1,2)	0.321	4.706	4.704
		(1,2)	(1,2)	0.351	4.855	4.814
		(1,3)	(1,2)	0.370	5.114	5.046
		(3,1)	(1,3)	0.310	6.021	5.733
		(2,1)	(1,3)	0.325	5.346	5.264
		(1,1)	(1,3)	0.356	4.677	4.675
		(1,2)	(1,3)	0.403	4.613	4.565
		(1,3)	(1,3)	0.443	4.747	4.654

Table A.4: Performance comparison of emergency stock resupply policies

π'	Total Stock	λ	Backorder Cost	η^*	Cost Reduction (%)	
					Optimal Policy	Approximate Policy
30	8	(1,1)	(1,1)	0.702	41.610	41.610
		(1,2)	(1,1)	0.700	41.694	41.694
		(1,3)	(1,1)	0.701	41.625	41.625
		(3,1)	(1,2)	0.757	41.248	41.224
		(2,1)	(1,2)	0.770	42.384	42.347
		(1,1)	(1,2)	0.819	44.517	44.450
		(1,2)	(1,2)	0.892	45.820	45.808
		(1,3)	(1,2)	0.940	48.169	47.490
		(3,1)	(1,3)	0.801	42.301	42.177
		(2,1)	(1,3)	0.833	44.193	44.059
		(1,1)	(1,3)	0.907	44.297	44.105
		(1,2)	(1,3)	1.023	50.177	50.177
		(1,3)	(1,3)	1.127	50.581	49.725
	12	(1,1)	(1,1)	0.289	43.835	43.835
		(1,2)	(1,1)	0.288	43.875	43.875
		(1,3)	(1,1)	0.289	43.729	43.729
		(3,1)	(1,2)	0.312	43.423	43.423
		(2,1)	(1,2)	0.317	44.659	44.659
		(1,1)	(1,2)	0.337	46.955	46.955
		(1,2)	(1,2)	0.369	48.148	47.677
		(1,3)	(1,2)	0.390	50.213	49.779
		(3,1)	(1,3)	0.330	44.545	44.545
		(2,1)	(1,3)	0.344	46.442	46.442
		(1,1)	(1,3)	0.373	46.856	46.447
		(1,2)	(1,3)	0.423	52.474	51.742
		(1,3)	(1,3)	0.465	53.011	52.291
50	8	(1,1)	(1,1)	0.702	24.621	24.621
		(1,2)	(1,1)	0.700	24.690	24.690
		(1,3)	(1,1)	0.701	24.662	24.662
		(3,1)	(1,2)	0.757	24.670	24.617
		(2,1)	(1,2)	0.770	25.404	25.343
		(1,1)	(1,2)	0.819	27.139	26.867
		(1,2)	(1,2)	0.892	28.827	28.389
		(1,3)	(1,2)	0.940	31.303	30.835
		(3,1)	(1,3)	0.801	25.607	25.607
		(2,1)	(1,3)	0.833	26.984	26.855
		(1,1)	(1,3)	0.907	27.670	27.276
		(1,2)	(1,3)	1.023	33.476	32.723
		(1,3)	(1,3)	1.127	34.826	34.278
	12	(1,1)	(1,1)	0.289	26.387	26.387
		(1,2)	(1,1)	0.288	26.431	26.431
		(1,3)	(1,1)	0.289	26.361	26.361
		(3,1)	(1,2)	0.312	26.435	26.410
		(2,1)	(1,2)	0.317	27.251	27.212
		(1,1)	(1,2)	0.337	29.159	29.005
		(1,2)	(1,2)	0.369	30.753	30.664
		(1,3)	(1,2)	0.390	33.115	32.216
		(3,1)	(1,3)	0.330	27.465	27.363
		(2,1)	(1,3)	0.344	28.874	28.780
		(1,1)	(1,3)	0.373	29.838	29.730
		(1,2)	(1,3)	0.423	35.602	35.291
		(1,3)	(1,3)	0.465	37.176	36.600

Appendix B

MATLAB Code

B.1 Solution Procedure for Stock Allocation

```
function [ V,AVG] = OptimalStockAllocation2Bases( X,N,Lambda,Mu,c,S )
lambda=Lambda./(sum(Lambda)+Mu);mu=Mu/(sum(Lambda)+Mu);
V=zeros(X,X,2);avg=zeros(1,N);
for n=1:N
    avg(n)=mean(mean(V(:,:,2)-V(:,:,1)));
    V(:,:,1)=V(:,:,2);
    for x1=0:X-1
        i1=x1+1;
        for x2=0:X-1
            i2=x2+1;
            et=lambda(1)*V(min(i1+1,X-1),i2,1)+lambda(2)*V(i1,min(i2+1,X-1),1);
            et1=mu*V(max(i1-1,1),i2,1)+et;
            et2=mu*V(i1,max(i2-1,1),1)+et;
            stcost=(c(1)*max(x1-S(1),0))+(c(2)*max(x2-S(2),0));
            V(i1,i2,2)=stcost+min(et1,et2);
        end
    end
end
V=V(:,:,2);
V=V-min(min(V));
AVG=avg(n);

function [ V,AVG,SAM,Sp ] = ValueFuncStockAllocation2Bases( X,N,Lambda,Mu,c,tS )
SAM=zeros(X,X);K=2;
SC(1,2)=floor(log(c(1)/c(2))/log(Lambda(2)/Mu));
S=zeros(1,K);g1=zeros(1,K);
for i=1:tS
    for j=1:K
        g1(j)=(c(j)*((Lambda(j)/Mu)^(S(j)+1))/(1-(Lambda(j)/Mu)))/(Lambda(j));
    end
    [~,SB]=max(g1);
    S(SB)=S(SB)+1;
end
Sp=S;
for x1=0:X-1
    i1=x1+1;
    s(1)=Sp(1)-x1;
    if s(1)>=0
        h1(1)=(c(1)*((Lambda(1)/Mu)^(s(1)+1))/(1-(Lambda(1)/Mu)))/(Lambda(1));
    end
```

```
    for x2=0:X-1
        i2=x2+1;
        s(2)=Sp(2)-x2;
        if s(2)>=0
            h1(2)=(c(2)*((Lambda(2)/Mu)^(s(2)+1))/(1-(Lambda(2)/Mu)))/(Lambda(2));
        end
        sam=compare2bases(h1(1),h1(2),s(1),s(2),Lambda(1),Lambda(2),SC(1,2),1,2);
        SAM(i1,i2)=sam;
    end
end
lambda=Lambda./(sum(Lambda)+Mu);mu=Mu/(sum(Lambda)+Mu);S=Sp;
V=zeros(X,X,2);avg=zeros(1,N);
for n=1:N
    avg(n)=mean(mean(V(:,:,2)-V(:,:,1)));
    V(:,:,1)=V(:,:,2);
    for x1=0:X-1
        i1=x1+1;
        for x2=0:X-1
            i2=x2+1;
            et0=lambda(1)*V(min(i1+1,X-1),i2,1)+lambda(2)*V(i1,min(i2+1,X-1),1);
            et(1)=mu*V(max(i1-1,1),i2,1)+et0;
            et(2)=mu*V(i1,max(i2-1,1),1)+et0;
            stcost=(c(1)*max(x1-S(1),0))+(c(2)*max(x2-S(2),0));
            V(i1,i2,2)=stcost+et(SAM(i1,i2));
        end
    end
end
V=V(:,:,2);
V=V-min(min(V));
AVG=avg(n);

function [ V,AVG,SAM,Sp ] = MyopicStockAllocation2Bases( X,N,Lambda,Mu,c,tS )
SAM=zeros(tS+1);
SC1=(log(c(2)/c(1))/log(Lambda(2)/Mu));
if SC1<0
    SC1=ceil(SC1);
else
    SC1=floor(SC1);
end
Sp(1)=0;Sp(2)=0;
for i=0:tS
    h1(1)=c(1)*(Lambda(1)/(Lambda(1)+Mu))^(Sp(1)+1);
    h1(2)=c(2)*(Lambda(2)/(Lambda(2)+Mu))^(Sp(2)+1);
    if h1(1)>h1(2)
        SAM(tS-Sp(1)+1,1:tS-Sp(2)+1)=1;
        Sp(1)=Sp(1)+1;
    elseif h1(1)==h1(2)
        if Lambda(1)>Lambda(2)
            SAM(tS-Sp(1)+1,1:tS-Sp(2)+1)=1;
            Sp(1)=Sp(1)+1;
        else
            SAM(1:tS-Sp(1)+1,tS-Sp(2)+1)=2;
            Sp(2)=Sp(2)+1;
        end
    else
        SAM(1:tS-Sp(1)+1,tS-Sp(2)+1)=2;
        Sp(2)=Sp(2)+1;
    end
    if sum(Sp)==tS
        break
    end
end
```

```
SAM=SAM(tS-Sp(1)+1:end,tS-Sp(2)+1:end);
SAM(1,1)=2;
SAM(Sp(1)+2:X,1:Sp(2)+SC1+1)=1;
SAM(1:X,Sp(2)+SC1+2:X)=2;
S=Sp;
lambda=Lambda./(sum(Lambda)+Mu);mu=Mu/(sum(Lambda)+Mu);
V=zeros(X,X,2);avg=zeros(1,N);
for n=1:N
    avg(n)=mean(mean(V(:,:,2)-V(:,:,1)));
    V(:,:,1)=V(:,:,2);
    for x1=0:X-1
        i1=x1+1;
        for x2=0:X-1
            i2=x2+1;
            et0=lambda(1)*V(min(i1+1,X-1),i2,1)+lambda(2)*V(i1,min(i2+1,X-1),1);
            et(1)=mu*V(max(i1-1,1),i2,1)+et0;
            et(2)=mu*V(i1,max(i2-1,1),1)+et0;
            stcost=(c(1)*max(x1-S(1),0))+(c(2)*max(x2-S(2),0));
            V(i1,i2,2)=stcost+et(SAM(i1,i2));
        end
    end
end
V=V(:,:,2);
V=V-min(min(V));
AVG=avg(n);

function [AVG,MAVG,S]=ValueFuncStockAllocationKBases(ran,N,rep,warmup,Lambda,Mu,c,tS)
K=length(Lambda);
SC=zeros(K-1,K);
for i=1:K-1
    for j=i+1:K
        SC(i,j)=floor(log(c(i)/c(j))/log(Lambda(j)/Mu));
    end
end
S=zeros(1,K);g1=zeros(1,K);
for i=1:tS
    for j=1:K
        g1(j)=(c(j)*((Lambda(j)/(Mu))^(S(j)+1))/(1-(Lambda(j)/(Mu))));%/((K)*Mu+Lambda(j));
    end
    [~,SB]=max(g1);
    S(SB)=S(SB)+1;
end
h1=zeros(K,tS);
for i=1:K
    for j=0:S(i)
        s=S(i)-j;
        h1(i,j+1)=(c(i)*((Lambda(i)/(Mu))^(s+1))/(1-(Lambda(i)/(Mu))));%/((K)*Mu+Lambda(i));
    end
end
lambda=Lambda/sum(Lambda);
t=tril(ones(K,K),0);
lambda=t*lambda';
lambda=lambda';
AVG=zeros(1,rep);
for l=1:rep
x=zeros(1,K);tcost=0;
for i=1:N
    r=ran(l,i);
    if r<(Mu/(sum(Lambda)+Mu))
        if max(x)>0
    B=find(IL<0,1,'last');
    if ~isempty(B)
```

```
        if B<K
        C=IL(B+1:end)-SC(B,B+1:end);
        BC=find(C<0,1,'last');
        if ~isempty(BC)
            sam=BC+B;
        else
            sam=B;
        end
        else
            sam=B;
        end
    else
        sam=1;h=h1(1,x(1)+1);
        for j=2:K
            if h1(j,x(j)+1)>h
                sam=j;
                h=h1(j,x(j)+1);
            elseif h1(j,x(j)+1)==h && Lambda(j)>Lambda(sam)
                sam=j;
                h=h1(j,x(j)+1);
            end
        end
    end
        g=zeros(1,K);
        g(sam)=1;
        x=max(x-g,0);
        end
    else
        r=(r-(Mu/(sum(Lambda)+Mu)))*((sum(Lambda)+Mu)/sum(Lambda));
        rm=lambda-r;
        dem=find(rm>0,1,'first');
        g=zeros(1,K);
        g(dem)=1;
        x=x+g;
    end
    IL=S-x;
    ILC=abs(min(0,IL));
    if i>warmup
    stcost=c*ILC';
    tcost=tcost+stcost;
    end
end
AVG(l)=tcost/(N-warmup);
end
MAVG=mean(AVG);

function [AVG,MAVG,S]=MyopicStockAllocationKBases(ran,N,rep,warmup,Lambda,Mu,c,tS)
K=length(Lambda);
SC=zeros(K-1,K);
for i=1:K-1
    for j=i+1:K
        SC(i,j)=floor(log(c(i)/c(j))/log(Lambda(j)/Mu));
    end
end
S=zeros(1,K);g1=zeros(1,K);
for i=1:tS
    for j=1:K
        g1(j)=c(j)*(Lambda(j)/(Lambda(j)+Mu))^(S(j)+1);
    end
    [~,SB]=max(g1);
    S(SB)=S(SB)+1;
end
```

```
h1=zeros(K,tS);
for i=1:K
    for j=0:S(i)
        s=S(i)-j;
        h1(i,j+1)=c(i)*(Lambda(i)/(Lambda(i)+Mu))^(s+1);
    end
end
lambda=Lambda/sum(Lambda);
t=tril(ones(K,K),0);
lambda=t*lambda';
lambda=lambda';
AVG=zeros(1,rep);
for l=1:rep
x=zeros(1,K);tcost=0;
for i=1:N
    r=ran(l,i);
    if r<(Mu/(sum(Lambda)+Mu))
        if max(x)>0
    B=find(IL<0,1,'last');
    if ~isempty(B)
        if B<K
        C=IL(B+1:end)-SC(B,B+1:end);
        BC=find(C<0,1,'last');
        if ~isempty(BC)
            sam=BC+B;
        else
            sam=B;
        end
        else
            sam=B;
        end
    else
        sam=1;h=h1(1,x(1)+1);
        for j=2:K
            if h1(j,x(j)+1)>h
                sam=j;
                h=h1(j,x(j)+1);
            elseif h1(j,x(j)+1)==h && Lambda(j)>Lambda(sam)
                sam=j;
                h=h1(j,x(j)+1);
            end
        end
    end
        g=zeros(1,K);
        g(sam)=1;
        x=max(x-g,0);
        end
    else
        r=(r-(Mu/(sum(Lambda)+Mu)))*((sum(Lambda)+Mu)/sum(Lambda));
        rm=lambda-r;
        dem=find(rm>0,1,'first');
        g=zeros(1,K);
        g(dem)=1;
        x=x+g;
    end
    IL=S-x;
    ILC=abs(min(0,IL));
    if i>warmup
    stcost=c*ILC';
    tcost=tcost+stcost;
    end
end
```

```
AVG(l)=tcost/N;
end
MAVG=mean(AVG);
```

B.2 Solution Procedure for Stock Reallocation

```
function [SAV,SAavg,V, avg ,lavg,SAM,SRAM,S]=OptimalStockReallocation2Bases(X,N,Lambda,Mu,c,ctr,trr,tS)
V=zeros(X,X,2);avg=zeros(1,N);
lambda1=Lambda./(sum(Lambda)+Mu+trr);trr1=trr/(sum(Lambda)+Mu+trr);mu1=Mu/(sum(Lambda)+Mu+trr);
[ SAV,SAavg,SAM,S ] = ValueFuncStockAllocation2Bases( X,N,Lambda,Mu,c,tS );
SRAM=zeros(X,X);
for n=1:N
    avg(n)=mean(mean(V(1:S(1),1:S(2),2)-V(1:S(1),1:S(2),1)));
    V(:,:,1)=V(:,:,2);
    for x1=0:X-1
        i1=x1+1;
        for x2=0:X-1
            i2=x2+1;
                etm1(1)=mu1*V(max(i1-1,1),i2,1);
                etm1(2)=mu1*V(i1,max(i2-1,1),1);
                sra0=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(V(i1,i2,1))+etm1(SAM(i1
                sra1=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(ctr+V(min(i1+1,X-1),max
                sra2=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(ctr+V(max(i1-1,1),min(i2
                    et=min(sra0,min(sra1,sra2));
                if x1>=S(1)&& x2>=S(2)
                    et=sra0;
                elseif x1>=S(1)
                    et=min(sra0,sra2);
                elseif x2>=S(2)
                    et=min(sra0,sra1);
                end
                if et==sra0
                    SRAM(i1,i2)=0;
                 elseif et==sra1
                    SRAM(i1,i2)=1;
                else
                    SRAM(i1,i2)=2;
                end
            stcost=((c(1)*max(x1-S(1),0))+(c(2)*max(x2-S(2),0)));
            V(i1,i2,2)=stcost+et;
        end
    end
end
AVG=V(:,:,2)-V(:,:,1);
V=V(:,:,2);
V=V-min(min(V));
lavg=avg(n);

function [ApprSAV, SAavg,V, avg ,lavg,SAM,SRAM]=ApprStockReallocation2Bases(X,N,Lambda,Mu,c,ctr,trr,tS)
lambda1=Lambda./(sum(Lambda)+Mu+trr);trr1=trr/(sum(Lambda)+Mu+trr);mu1=Mu/(sum(Lambda)+Mu+trr);
[ SAV,SAavg,SAM,S ] = ValueFuncStockAllocation2Bases( X,N,Lambda,Mu,c,tS );
V=SAV;
V(1:S(1)+1,1:S(2)+1)=hsValFunAppr(Lambda,Mu,c,S,SAM,ctr,trr);
ApprSAV=V;
SRAM=zeros(X,X);SC=S(2)-floor(log(c(1)/c(2))/log(Lambda(2)/Mu));
for x1=0:X-1
    i1=x1+1;
    for x2=0:X-1
        i2=x2+1;
        etm1(1)=mu1*V(max(i1-1,1),i2,1);
```

```
        etm1(2)=mu1*V(i1,max(i2-1,1),1);
        sra0=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(V(i1,i2,1))+etm1(SAM(i1,i2));
        sra1=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(ctr+V(min(i1+1,X-1),max(i2-1,1)
        sra2=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(ctr+V(max(i1-1,1),min(i2+1,X-1)
        et=min(sra0,min(sra1,sra2));
        if x1>=S(1)&& x2>=SC
            et=sra0;
        end
        if et==sra0
            SRAM(i1,i2)=0;
        elseif et==sra1
            SRAM(i1,i2)=1;
            if SAM(i1,i2)==1
                SRAM(i1,i2)=0;
            end
        else
            SRAM(i1,i2)=2;
            if SAM(i1,i2)==2
                SRAM(i1,i2)=0;
            end
        end
    end
end
V=zeros(X,X,2);avg=zeros(1,N);
for n=1:N
    avg(n)=mean(mean(V(1:S(1),1:S(2),2)-V(1:S(1),1:S(2),1)));
    V(:,:,1)=V(:,:,2);
    for x1=0:X-1
        i1=x1+1;
        for x2=0:X-1
            i2=x2+1;
            etm1(1)=mu1*V(max(i1-1,1),i2,1);
            etm1(2)=mu1*V(i1,max(i2-1,1),1);
            sra0=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(V(i1,i2,1))+etm1(SAM(i1,i2))
            sra1=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(ctr+V(min(i1+1,X-1),max(i2-
            sra2=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+trr1*(ctr+V(max(i1-1,1),min(i2+1,X
            if SRAM(i1,i2)==0
                et=sra0;
            elseif SRAM(i1,i2)==1
                et=sra1;
            else
                et=sra2;
            end
            stcost=((c(1)*max(x1-S(1),0))+(c(2)*max(x2-S(2),0)));
            V(i1,i2,2)=stcost+et;
        end
    end
end
V=V(:,:,2);
V=V-min(min(V));
lavg=avg(n);

function h=hsValFunAppr(Lambda,Mu,c,S,g,ctr,trr)
[v1,~,~]=SinQueueValFunWTRR(100,10000,Lambda(1),Mu,trr,c(1),ctr,S(1));
[v2,~,~]=SinQueueValFunWTRR(100,10000,Lambda(2),Mu,trr,c(2),ctr,S(2));
[v,~]=SinQValFun(100,10000,Mu,sum(Lambda),c(1),0,sum(S));
for i=1:S(1)+1
    for j=1:S(2)+1
        if g(i,j)==1
           h(i,j)=v(i+j-1)+v1(i);
        elseif g(i,j)==2
            h(i,j)=v(i+j-1)+v2(j);
```

```
        end
    end
end

function [AVG,SRA,MAVG,S]=StockReallocationKBases(sra,ran,N,rep,warmup,Lambda,Mu,c,ctr,trr,tS)
K=length(Lambda);SRAcost=0;
SC=zeros(K-1,K);
for i=1:K-1
    for j=i+1:K
        SC(i,j)=floor(log(c(i)/c(j))/log(Lambda(j)/Mu));
    end
end
S=zeros(1,K);g1=zeros(1,K);
for i=1:tS
    for j=1:K
        g1(j)=(c(j)*((Lambda(j)/(Mu))^(S(j)+1))/(1-(Lambda(j)/(Mu))))/((K-2)*Mu+Lambda(j));
    end
    [~,SB]=max(g1);
    S(SB)=S(SB)+1;
end
h1=zeros(K,tS);
for i=1:K
    for j=0:S(i)
        s=S(i)-j;
        h1(i,j+1)=(c(i)*((Lambda(i)/(Mu))^(s+1))/(1-(Lambda(i)/(Mu))))/((K-2)*Mu+Lambda(i));
    end
end
for i=1:K
    [~,~,SR(i,:)]=SinQueueValFunWTRR(100,10000,Lambda(i),Mu,trr,c(i),ctr,S(i));
end
mu1=Mu/(trr+sum(Lambda)+Mu);
trr1=trr/(trr+sum(Lambda)+Mu);
lambda=Lambda/(trr+sum(Lambda)+Mu);
t=tril(ones(K,K),0);
lambda=t*lambda';
lambda=lambda';
lambda=lambda+mu1+trr1;
AVG=zeros(1,rep);
for l=1:rep
    x=zeros(1,K);tcost=0;
    for i=1:N
        r=ran(l,i);
        if r<trr1
            if sra
            if max(x)>0
                B=find(IL<0,1,'last');
                if ~isempty(B)
                    if B<K
                        C=IL(B+1:end)-SC(B,B+1:end);
                        BC=find(C<0,1,'last');
                        if ~isempty(BC)
                            sam=BC+B;
                        else
                            sam=B;
                        end
                    else
                        sam=B;
                    end
                else
                    sam=1;h=h1(1,x(1)+1);
                    for j=2:K
                        if h1(j,x(j)+1)>h
```

```
                sam=j;
                h=h1(j,x(j)+1);
            elseif h1(j,x(j)+1)==h && Lambda(j)>Lambda(sam)
                sam=j;
                h=h1(j,x(j)+1);
            end
        end
    end
    if SR(sam,x(sam)+1)==1
        [SAP,~]=SASequence(x+1,Lambda,Mu,c,S);
        rsam=SAP(length(SAP));
         if x(sam)>0 && IL(rsam)>0  && SR(rsam,x(rsam)+2)==0
            x(sam)=x(sam)-1;
            x(rsam)=x(rsam)+1;
            if i>warmup
                SRAcost=SRAcost+ctr;
            end
        end
    end
end
end
elseif r<trr1+mu1
    if max(x)>0
        B=find(IL<0,1,'last');
        if ~isempty(B)
            if B<K
                C=IL(B+1:end)-SC(B,B+1:end);
                BC=find(C<0,1,'last');
                if ~isempty(BC)
                    sam=BC+B;
                else
                    sam=B;
                end
            else
                sam=B;
            end
        else
            sam=1;h=h1(1,x(1)+1);
            for j=2:K
                if h1(j,x(j)+1)>h
                    sam=j;
                    h=h1(j,x(j)+1);
                elseif h1(j,x(j)+1)==h && Lambda(j)>Lambda(sam)
                    sam=j;
                    h=h1(j,x(j)+1);
                end
            end
        end
        g=zeros(1,K);
        g(sam)=1;
        x=max(x-g,0);
    end
else
    rm=lambda-r;
    dem=find(rm>0,1,'first');
    g=zeros(1,K);
    g(dem)=1;
    x=x+g;
end
IL=S-x;
ILC=abs(min(0,IL));
if i>warmup
```

```
            stcost=c*ILC';
            tcost=tcost+stcost;
        end
    end
    AVG(l)=tcost/(N-warmup);
    SRA(l)=SRAcost/(N-warmup);
end
MAVG=mean(AVG+SRA);
```

B.3 Solution Procedure for Emergency Resupply

```
function [SAV,SAavg, avg ,lavg,SAM,ERM,AVG,S]=OptimalEmergencyResupply2Bases(X,N,Lambda,Mu,c,cer,tS)
V=zeros(X,X,2);avg=zeros(1,N);
lambda1=Lambda./(sum(Lambda)+Mu);mu1=Mu/(sum(Lambda)+Mu);
[ SAV,SAavg,SAM,S ] = ValueFuncStockAllocation2Bases( X,N,Lambda,Mu,c,tS );
ERM=zeros(X,X);
for n=1:N
    avg(n)=mean(mean(V(1:S(1),1:S(2),2)-V(1:S(1),1:S(2),1)));
    V(:,:,1)=V(:,:,2);
    for x1=0:X-1
        i1=x1+1;
        for x2=0:X-1
            i2=x2+1;
                etm1(1)=mu1*V(max(i1-1,1),i2,1);
                etm1(2)=mu1*V(i1,max(i2-1,1),1);
                sra0=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+etm1(SAM(i1,i2));
                sra1=lambda1(1)*(cer+V(i1,i2,1))+lambda1(2)*V(i1,min(i2+1,X-1),1)+etm1(SAM(i1,i2));
                sra2=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*(cer+V(i1,i2,1))+etm1(SAM(i1,i2));
                sra3=lambda1(1)*(cer+V(i1,i2,1))+lambda1(2)*(cer+V(i1,i2,1))+etm1(SAM(i1,i2));
                if x1>=S(1)&& x2>=S(2)
                    et=min(sra0,min(sra1,min(sra2,sra3)));
                if et==sra1
                    ERM(i1,i2)=1;
                elseif et==sra2
                    ERM(i1,i2)=2;
                elseif et==sra3
                    ERM(i1,i2)=3;
                end
                else
                    et=sra0;
                end
            stcost=((c(1)*max(x1-S(1),0))+(c(2)*max(x2-S(2),0)));
            V(i1,i2,2)=stcost+et;
        end
    end
end
AVG=V(:,:,2)-V(:,:,1);
V=V(:,:,2);
V=V-min(min(V));
lavg=avg(n);

function h=KlimovValFunKBases(x,Lambda,Mu,c)
K=length(Lambda);
D=zeros(K,K);Dp=zeros(1,K);
Dp(1)=c(1)/(Mu-sum(Lambda));D(1,1:K)=Dp(1);D(1:K,1)=Dp(1);
for i=2:K
    Dp(i)=(c(i)+sum(Lambda(1:i-1).*Dp(1:i-1)))/(Mu-sum(Lambda(i:K)));
    D(i,i:K)=Dp(i);D(i:K,i)=Dp(i);
```

```
end
h=(1/2)*(x*(D*x'))+(1/2)*(Dp*x');
h=h*(sum(Lambda)+Mu);

function hmod=ModKlimovValFunKBases(x,horig,Lambda,Mu,c)
K=length(Lambda);xp=zeros(1,K);xp(1)=sum(x);
Delta=KlimovValFunKBases(x,Lambda,Mu,c)-KlimovValFunKBases(xp,Lambda,Mu,c);
hmod=horig(xp(1)+1)+Delta;

function [V, avg ,lavg,SAM,ERM,AVG]=ApprEmergencyResupply2Bases(X,N,Lambda,Mu,c,cer,S,SAM)
avg=zeros(1,N);
lambda1=Lambda./(sum(Lambda)+Mu);mu1=Mu/(sum(Lambda)+Mu);
ERM=zeros(X,X);
[h]=h0ValFunApprRK2(X,Lambda,Mu,c,cer,S);
V=zeros(X,X);V(S(1)+1:X,S(2)+1:X)=h(1:X-S(1),1:X-S(2));V=V*(sum(Lambda)+Mu);
for x1=0:X-1
    i1=x1+1;
    for x2=0:X-1
        i2=x2+1;
        etm1(1)=mu1*V(max(i1-1,1),i2,1);
        etm1(2)=mu1*V(i1,max(i2-1,1),1);
        sra0=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+etm1(SAM(i1,i2));
                sra1=lambda1(1)*(cer+V(i1,i2,1))+lambda1(2)*V(i1,min(i2+1,X-1),1)+etm1(SAM(i1,i2));
                sra2=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*(cer+V(i1,i2,1))+etm1(SAM(i1,i2));
                sra3=lambda1(1)*(cer+V(i1,i2,1))+lambda1(2)*(cer+V(i1,i2,1))+etm1(SAM(i1,i2));
                if x1>=S(1)&& x2>=S(2)
                    et=min(sra0,min(sra1,min(sra2,sra3)));
                if et==sra1
                    ERM(i1,i2)=1;
                elseif et==sra2
                    ERM(i1,i2)=2;
                elseif et==sra3
                    ERM(i1,i2)=3;
                end
                end
    end
end
V=zeros(X,X,2);
for n=1:N
    avg(n)=mean(mean(V(1:S(1),1:S(2),2)-V(1:S(1),1:S(2),1)));
    V(:,:,1)=V(:,:,2);
    for x1=0:X-1
        i1=x1+1;
        for x2=0:X-1
            i2=x2+1;
            etm1(1)=mu1*V(max(i1-1,1),i2,1);
            etm1(2)=mu1*V(i1,max(i2-1,1),1);
            if ERM(i1,i2)==1
                et=lambda1(1)*(cer+V(i1,i2,1))+lambda1(2)*V(i1,min(i2+1,X-1),1)+etm1(SAM(i1,i2));
            elseif ERM(i1,i2)==2
                et=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*(cer+V(i1,i2,1))+etm1(SAM(i1,i2));
            elseif ERM(i1,i2)==3
                et=lambda1(1)*(cer+V(i1,i2,1))+lambda1(2)*(cer+V(i1,i2,1))+etm1(SAM(i1,i2));
            else
                et=lambda1(1)*V(min(i1+1,X-1),i2,1)+lambda1(2)*V(i1,min(i2+1,X-1),1)+etm1(SAM(i1,i2));
            end
            stcost=((c(1)*max(x1-S(1),0))+(c(2)*max(x2-S(2),0)));
            V(i1,i2,2)=stcost+et;
        end
    end
end
```

```
AVG=V(:,:,2)-V(:,:,1);
V=V(:,:,2);
V=V-min(min(V));
lavg=avg(n);

function [h]=h0ValFunApprRK2(X,Lambda,Mu,c,cer,S)
delta=0;
[v,~,~]=CapSinQueueRelValFun(X+sum(S)+2,sum(Lambda),Mu,c(1),cer,sum(S));
[~,h]=KlimovValFun2(X+1,Lambda,Mu,c);
d=zeros(length(h(:,1)),length(h(1,:)));
for i=1:length(h(:,1))-1
    for j=2:length(h(1,:))
        d(i,j)=h(i,j)-h(i+1,j-1)-delta;
    end
end
h(:,1)=v(sum(S)+1:length(v));
for j=2:length(h(1,:))
    for i=1:length(h(:,1))-1
        h(i,j)=h(i+1,j-1)+d(i,j);
    end
end

function [MAVG,AVG,ERAVG,S]=EmergencyResupplyKBases(er,ran,N,rep,warmup,Lambda,Mu,c,cer,tS)
K=length(Lambda);
SC=zeros(K-1,K);
for i=1:K-1
    for j=i+1:K
        SC(i,j)=floor(log(c(i)/c(j))/log(Lambda(j)/Mu));
    end
end
S=zeros(1,K);g1=zeros(1,K);
for i=1:tS
    for j=1:K
        g1(j)=(c(j)*((Lambda(j)/(Mu))^(S(j)+1))/(1-(Lambda(j)/(Mu))));%/((K)*Mu+Lambda(j));
    end
    [~,SB]=max(g1);
    S(SB)=S(SB)+1;
end
h1=zeros(K,tS);
for i=1:K
    for j=0:S(i)
        s=S(i)-j;
        h1(i,j+1)=(c(i)*((Lambda(i)/(Mu))^(s+1))/(1-(Lambda(i)/(Mu))));%/((K)*Mu+Lambda(i));
    end
end
[horig,~,~]=CapSinQueueRelValFun(100,sum(Lambda),Mu,c(1),cer,tS);
horig=horig(tS+1:end);
lambda=Lambda/sum(Lambda);
t=tril(ones(K,K),0);
lambda=t*lambda';
lambda=lambda';
AVG=zeros(1,rep);ERAVG=AVG;
for l=1:rep
    x=zeros(1,K);tcost=0;ERCost=0;
    for i=1:N
        r=ran(l,i);
        if r<(Mu/(sum(Lambda)+Mu))
            if max(x)>0
                B=find(IL<0,1,'last');
                if ~isempty(B)
                    if B<K
```

```
                    C=IL(B+1:end)-SC(B,B+1:end);
                    BC=find(C<0,1,'last');
                    if ~isempty(BC)
                        sam=BC+B;
                    else
                        sam=B;
                    end
                else
                    sam=B;
                end
            else
                sam=1;h=h1(1,x(1)+1);
                for j=2:K
                    if h1(j,x(j)+1)>h
                        sam=j;
                        h=h1(j,x(j)+1);
                    elseif h1(j,x(j)+1)==h && Lambda(j)>Lambda(sam)
                        sam=j;
                        h=h1(j,x(j)+1);
                    end
                end
            end
            g=zeros(1,K);
            g(sam)=1;
            x=max(x-g,0);
        end
    else
        r=(r-(Mu/(sum(Lambda)+Mu)))*((sum(Lambda)+Mu)/sum(Lambda));
        rm=lambda-r;
        dem=find(rm>0,1,'first');
        g=zeros(1,K);
        g(dem)=1;
        if ~er
            x=x+g;
        elseif min(x>=S)==1
            xp=x-S;
            delta=ModKlimovValFunKBases(xp+g,horig,Lambda,Mu,c)-ModKlimovValFunKBases(xp,horig,Lambda,Mu,c);
            if delta>cer
                if i>warmup
                    ERCost=ERCost+cer;
                end
            else
                x=x+g;
            end
        else
            x=x+g;
        end
    end
    IL=S-x;
    ILC=abs(min(0,IL));
    if i>warmup
        stcost=c*ILC';
        tcost=tcost+stcost;
    end
end
AVG(l)=tcost/(N-warmup);
ERAVG(l)=ERCost/(N-warmup);
end
MAVG=mean(AVG+ERAVG);
```

References

[1] I. Ferguson, "Assessing total MRO spend," in *Central and Eastern European Airline E & M And Supply Chain Conference*, 2009.

[2] C. C. Sherbrooke, "Metric: A multi-echelon technique for recoverable item control," *Operations Research*, vol. 16, no. 1, pp. 122–141, 1968.

[3] J. A. Muckstadt, *Analysis and Algorithms for Service Parts Supply Chains*. Springer, 2005.

[4] K. Michaels and J. Murby, "Air transport demand growth at an inflexion point," in *24th Annual Commercial Aviation Industry Suppliers Conference*, 2010.

[5] M. A. Cohen and J. Wille, "Implications for service parts management in the rapidly changing aviation MRO market," *Working Paper*, 2006.

[6] J. M. Berger, "Global MRO market: forcast and trends," in *Aircraft Maintenance Russia and CIS*, 2013.

[7] K. L. Cheung and W. H. Hausman, "Joint determination of preventive maintenance and safety stocks in an unreliable production environment," *Naval Research Logistics*, vol. 44, pp. 257–272, Apr. 1997.

[8] T. K. Das and S. Sarkar, "Optimal preventive maintenance in a production inventory system," *IIE Transactions*, vol. 31, pp. 537–551, 1999.

[9] K. S. de Smidt-Destombes, M. C. van der Heijden, and A. van Harten, "Joint optimisation of spare part inventory, maintenance frequency and repair capacity for k-out-of-n systems," *International Journal of Production Economics*, vol. 118, no. 1, pp. 260–268, 2009.

[10] A. H. Elwany and N. Z. Gebraeel, "Sensor-driven prognostic models for equipment replacement and spare parts inventory," *IIE - Transactions*, vol. 40, no. 7, p. 629, 2008.

[11] M. A. Ilgin and S. Tunali, "Joint optimization of spare parts inventory and maintenance policies using genetic algorithms," *The International Journal of Advanced Manufacturing Technology*, vol. 34, pp. 594–604, May 2006.

[12] S.-J. Joo, "Scheduling preventive maintenance for modular designed components: A dynamic approach," *European Journal of Operational Research*, vol. 192, pp. 512–520, Jan. 2009.

[13] Y.-H. Cheng and H.-L. Tsao, "Rolling stock maintenance strategy selection, spares parts' estimation, and replacements' interval calculation," *International Journal of Production Economics*, vol. 128, pp. 404–412, Nov. 2010.

[14] G. Lanza, S. Niggeschmidt, and P. Werner, "Optimization of preventive maintenance and spare part provision for machine tools based on variable operational conditions," *CIRP Annals - Manufacturing Technology*, vol. 58, no. 1, pp. 429–432, 2009.

[15] T. S. Vaughan, “Failure replacement and preventive maintenance spare parts ordering policy,” *European Journal of Operational Research*, vol. 161, pp. 183–190, Feb. 2005.

[16] A. B. M. Zohrul Kabir and A. S. Al-Olayan, “A stocking policy for spare part provisioning under age based preventive replacement,” *European Journal of Operational Research*, vol. 90, no. 1, pp. 171–181, 1996.

[17] I. J. Adan, A. Sleptchenko, and G. Van Houtum, “Reducing costs of spare parts supply systems via static priorities,” *Asia-Pacific Journal of Operational Research*, vol. 26, no. 04, pp. 559–585, 2009.

[18] B. Kat and Z. M. Avsar, “Using aggregate fill rate for dynamic scheduling of multi-class systems,” *Annals of Operations Research*, vol. 182, no. 1, pp. 87–117, 2011.

[19] A. Khatab, N. Nahas, and M. Nourelfath, “Availability of k-out-of-N:G systems with non-identical components subject to repair priorities,” *Reliability Engineering & System Safety*, vol. 94, pp. 142–151, Feb. 2009.

[20] K. N. F. Leung, Y. L. Zhang, and K. K. Lai, “A bivariate optimal replacement policy for a cold standby repairable system with repair priority,” *Naval Research Logistics*, vol. 57, no. 2, pp. 149–158, 2010.

[21] Y. L. Zhang and G. J. Wang, “A geometric process repair model for a repairable cold standby system with priority in use and repair,” *Reliability Engineering & System Safety*, vol. 94, pp. 1782–1787, Nov. 2009.

[22] D. F. Pyke, "Priority repair and dispatch policies for reparable-item logistics systems," *Naval Research Logistics*, vol. 37, pp. 1–30, Feb. 1990.

[23] A. Sleptchenko, M. C. van der Heijden, and A. van Harten, "Using repair priorities to reduce stock investment in spare part networks," *European Journal of Operational Research*, vol. 163, no. 3, pp. 733–750, 2005.

[24] A. Diaz and M. C. Fu, "Models for multi-echelon repairable item inventory systems with limited repair capacity," *European Journal of Operational Research*, vol. 97, pp. 480–492, Mar. 1997.

[25] B. Jung, B. Sun, J. Kim, and S. Ahn, "Modeling lateral transshipments in multiechelon repairable-item inventory systems with finite repair channels," *Computers & Operations Research*, vol. 30, no. 9, pp. 1401–1417, 2003.

[26] J. A. Rappold and B. D. Van Roo, "Designing multi-echelon service parts networks with finite repair capacity," *European Journal of Operational Research*, vol. 199, pp. 781–792, Dec. 2009.

[27] A. Sleptchenko, M. C. van der Heijden, and A. van Harten, "Effects of finite repair capacity in multi-echelon, multi-indenture service part supply systems," *International Journal of Production Economics*, vol. 79, pp. 209–230, Oct. 2002.

[28] L. A. Baxter, "On the theory of cannibalization," *Journal of Mathematical Analysis and Applications*, vol. 136, pp. 290–297, Nov. 1988.

[29] W. W. Fisher, "Markov process modelling of a maintenance system with

spares, repair, cannibalization and manpower constraints," *Mathematical and Computer Modelling*, vol. 13, no. 7, pp. 119–125, 1990.

[30] S. Salman, C. R. Cassady, E. A. Pohl, and S. W. Ormon, "Evaluating the impact of cannibalization on fleet performance," *Quality and Reliability Engineering International*, vol. 23, pp. 445–457, June 2007.

[31] C. S. Sung and S. H. Kim, "Analysis of a multi-part spares inventory system subject to ambiguous fault isolation," *Journal of the Operational Research Society*, vol. 52, pp. 418–432, Apr. 2001.

[32] R. Anupindi, Y. Bassok, and E. Zemel, "A general framework for the study of decentralized distribution systems," *Manufacturing and Service Oprtaions Management*, vol. 3, pp. 349–368, Jan. 2001.

[33] Z. M. Avsar and M. Baykal-Gursoy, "Inventory control under substitutable demand: A stochastic game application," *Naval Research Logistics*, vol. 49, pp. 359–375, June 2002.

[34] B. C. Hartman, M. Dror, and M. Shaked, "Cores of inventory centralization games," *Games and Economic Behavior*, vol. 31, pp. 26–49, Apr. 2000.

[35] J. Kilpi, J. Toyli, and A. Vepsalainen, "Cooperative strategies for the availability service of repairable aircraft components," *International Journal of Production Economics*, vol. 117, no. 2, pp. 360–370, 2009.

[36] P. M. Reyes, "A game theory approach for solving the transshipment problem: a supply chain management strategy teaching tool," *Supply Chain Management: An International Journal*, vol. 11, no. 4, pp. 288–293, 2006.

[37] N. Rudi, S. Kapur, and D. F. Pyke, "A two-location inventory model with transshipment and local decision making," *Management Science*, vol. 47, no. 12, pp. 1668–1680, 2001.

[38] H. Wong, D. V. Oudheusden, and D. Cattrysse, "Cost allocation in spare parts inventory pooling," *Transportation Research Part E: Logistics and Transportation Review*, vol. 43, no. 4, pp. 370–386, 2007.

[39] J. E. Boylan and A. A. Syntetos, "Spare parts management: a review of forecasting research and extensions," *IMA Journal of Management Mathematics*, vol. 21, pp. 227 –237, July 2010.

[40] A. H. C. Eaves and B. G. Kingsman, "Forecasting for the ordering and stock-holding of spare parts," *Journal of the Operational Research Society*, vol. 55, pp. 431–437, Apr. 2004.

[41] A. A. Ghobbar and C. H. Friend, "Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model," *Computers & Operations Research*, vol. 30, pp. 2097–2114, Dec. 2003.

[42] E. Porras and R. Dekker, "An inventory control system for spare parts at a refinery: An empirical comparison of different re-order point methods," *European Journal of Operational Research*, vol. 184, no. 1, pp. 101–132, 2008.

[43] E. A. Shale, J. E. Boylan, and F. R. Johnston, "Forecasting for intermittent demand: the estimation of an unbiased average," *Journal of the Operational Research Society*, vol. 57, no. 5, pp. 588–592, 2006.

[44] A. Syntetos, M. Keyes, and M. Babai, "Demand categorisation in a european spare parts logistics network," *International Journal of Operations & Production Management*, vol. 29, no. 3, pp. 292–316, 2009.

[45] T. R. Willemain, C. N. Smart, and H. F. Schwarz, "A new approach to forecasting intermittent demand for service parts inventories," *International Journal of Forecasting*, vol. 20, no. 3, pp. 375–387, 2004.

[46] J. K. Cochran and T. P. Lewis, "Computing small-fleet aircraft availabilities including redundancy and spares," *Computers & Operations Research*, vol. 29, no. 5, pp. 529–540, 2002.

[47] H. C. Lau, H. Song, C. T. See, and S. Y. Cheng, "Evaluation of time-varying availability in multi-echelon spare parts systems with passivation," *European Journal of Operational Research*, vol. 170, pp. 91–105, Apr. 2006.

[48] Z. Hua, J. Yang, F. Huang, and X. Xu, "A static-dynamic strategy for spare part inventory systems with nonstationary stochastic demand," *Journal of Operational Research Society*, vol. 60, no. 9, pp. 1254–1263, 2008.

[49] T. Jin and H. Liao, "Spare parts inventory control considering stochastic growth of an installed base," *Computers & Industrial Engineering*, vol. 56, no. 1, pp. 452–460, 2009.

[50] H. Liao, P. Wang, T. Jin, and S. Repaka, "Spare parts management considering new sales," in *Reliability and Maintainability Symposium*, (Los Alamitos, CA, USA), pp. 502–507, 2008.

[51] R. M. Hill, M. Omar, and D. K. Smith, "Stock replenishment policies for

a stochastic exponentially-declining demand process," *European Journal of Operational Research*, vol. 116, pp. 374–388, July 1999.

[52] P. P. Datta and R. Roy, "Cost modelling techniques for availability type service support contracts: A literature review and empirical study," *CIRP Journal of Manufacturing Science and Technology*, vol. 3, no. 2, pp. 142–157, 2010.

[53] P. Hypko, M. Tilebein, and R. Gleich, "Clarifying the concept of performance-based contracting in manufacturing industries: A research synthesis," *Journal of Service Management*, vol. 21, no. 5, pp. 625–655, 2010.

[54] S.-H. Kim, M. A. Cohen, and S. Netessine, "Performance contracting in after-sales service supply chains," *Management Science*, vol. 53, pp. 1843–1858, Dec. 2007.

[55] D. Nowicki, U. D. Kumar, H. J. Steudel, and D. Verma, "Spares provisioning under performance-based logistics contract: profit-centric approach," *Journal of the Operational Research Society*, vol. 59, pp. 342–352, Nov. 2006.

[56] W. Wang, "A model for maintenance service contract design, negotiation and optimization," *European Journal of Operational Research*, vol. 201, pp. 239–246, Feb. 2010.

[57] M. Pourakbar and R. Dekker, "Customer differentiated end-of-life inventory problem," *European Journal of Operational Research*, vol. 222, no. 1, pp. 44 – 53, 2012.

[58] R. H. Teunter and W. K. Klein Haneveld, "Inventory control of service parts in the final phase," *European Journal of Operational Research*, vol. 137, pp. 497–511, Mar. 2002.

[59] J. P. J. van Kooten and T. Tan, "The final order problem for repairable spare parts under condemnation," *Journal of Operational Research Society*, vol. 60, no. 10, pp. 1449–1461, 2008.

[60] M. Braglia, A. Grassi, and R. Montanari, "Multi-attribute classification method for spare parts inventory management," *Journal of Quality in Maintenance Engineering*, vol. 10, no. 1, pp. 55–65, 2004.

[61] S. Cavalieri, M. Garetti, M. Macchi, and R. Pinto, "A decision-making framework for managing maintenance spare parts," *Production Planning & Control: The Management of Operations*, vol. 19, no. 4, pp. 379–396, 2008.

[62] Y. Huang, D. Sun, G. Xing, and H. Chang, "Criticality evaluation for spare parts based on BP neural network," in *Artificial Intelligence and Computational Intelligence (AICI), 2010 International Conference on*, vol. 1, pp. 204–206, 2010.

[63] C. R. Schultz, "Spare parts inventory and cycle time reduction," *International Journal of Production Research*, vol. 42, no. 4, p. 759, 2004.

[64] C. C. Sherbrooke, "Vari-metric: improved approximations for multi-indenture, multi-echelon availability models," *Operations Research*, vol. 34, no. 2, pp. 311–319, 1986.

[65] A. Gupta and S. C. Albright, “Steady-state approximations for a multi-echelon multi-indentured repairable-item inventory system,” *European Journal of Operational Research*, vol. 62, no. 3, pp. 340–353, 1992.

[66] W. Zijm and Z. M. Avsar, “Capacitated two-indenture models for repairable item systems,” *International Journal of Production Economics*, vol. 81-82, pp. 573–588, Jan. 2003.

[67] Y. Song and H. C. Lau, “A periodic-review inventory model with application to the continuous-review obsolescence problem,” *European Journal of Operational Research*, vol. 159, pp. 110–120, Nov. 2004.

[68] W. Van Jaarsveld and R. Dekker, “Estimating obsolescence risk from demand data to enhance inventory controla case study,” *International Journal of Production Economics*, vol. 133, no. 1, pp. 423–431, 2011.

[69] R. M. Simon, “Stationary properties of a two-echelon inventory model for low demand items,” *Operations Research*, vol. 19, no. 3, pp. 761–773, 1971.

[70] S. C. Graves, “A Multi-Echelon inventory model for a repairable item with One-for-One replenishment,” *Management Science*, vol. 31, no. 10, pp. 1247–1256, 1985.

[71] S. Axsater, “Simple solution procedures for a class of two-echelon inventory problems,” *Operations Research*, vol. 38, no. 1, pp. 64–69, 1990.

[72] J. Taylor and R. R. P. Jackson, “An application of the birth and death process to the provision of spare machines,” *Operations Research*, vol. 5, no. 4, pp. 95–108, 1954.

[73] F. J. Toft and H. Boothroyd, "A queueing model for spare coal faces," *Operations Research*, vol. 10, no. 4, pp. 245–251, 1959.

[74] N. M. Mirasol, "A queuing approach to logistics systems," *Operations Research*, vol. 12, no. 5, pp. 707–724, 1964.

[75] D. Gross, H. D. Kahn, and J. D. Marsh, "Queueing models for spares provisioning," *Naval Research Logistics Quarterly*, vol. 24, no. 4, pp. 521–536, 1977.

[76] D. Gross, L. C. Kioussis, and D. R. Miller, "A network decomposition approach for approximating the steady-state behavior of markovian multi-echelon reparable item inventory systems," *Management Science*, vol. 33, no. 11, pp. 1453–1468, 1987.

[77] D. Gross, B. Gu, and R. M. Soland, "Iterative solution methods for obtaining the steady-state probability distributions of markovian multi-echelon repairable item inventory systems," *Computers & Operations Research*, vol. 20, no. 8, pp. 817–828, 1993.

[78] S. C. Albright and A. Soni, "Markovian multiechelon repairable inventory system," *Naval Research Logistics*, vol. 35, no. 1, pp. 49–61, 1988.

[79] Y. Wang, M. A. Cohen, and Y. Zheng, "A two-echelon repairable inventory system with stocking-center-dependent depot replenishment lead times," *Management Science*, vol. 46, no. 11, pp. 1441–1453, 2000.

[80] D. Caglar, C. Li, and D. Simchi-Levi, "Two-echelon spare parts inventory

system subject to a service constraint," *IIE Transactions*, vol. 36, no. 7, pp. 655–666, 2004.

[81] B. L. Miller, "Dispatching from depot repair in a recoverable item inventory system: On the optimality of a heuristic rule," *Management Science*, vol. 21, pp. 316–325, Nov. 1974.

[82] A. Federgruen and P. Zipkin, "Approximations of dynamic, multilocation production and inventory problems," *Management Science*, vol. 30, pp. 69–84, Jan. 1984.

[83] M. D. Buyukkurt and M. Parlar, "A comparison of allocation policies in a two-echelon repairable-item inventory model," *International Journal of Production Economics*, vol. 29, pp. 291–302, May 1993.

[84] K. E. Caggiano, J. A. Muckstadt, and J. A. Rappold, "Integrated real-time capacity and inventory allocation for reparable service parts in a two-echelon supply system," *Manufacturing and Service Operations Management*, vol. 8, pp. 292–319, Jan. 2006.

[85] F. Barahona, P. Chowdhary, M. Ettl, P. Huang, T. Kimbrel, L. Ladanyi, Y. Lee, B. Schieber, K. Sourirajan, M. Sviridenko, and G. Swirszcz, "Inventory allocation and transportation scheduling for logistics of network-centric military operations," *IBM Journal of Research and Development*, vol. 51, no. 3.4, pp. 391–407, 2007.

[86] H. Tiemessen and G. van Houtum, "Reducing costs of repairable inventory supply systems via dynamic scheduling," *International Journal of Production Economics*, vol. 143, pp. 478–488, Aug. 2012.

[87] H. L. Lee, "A multi-echelon inventory model for repairable items with emergency lateral transshipments," *Management Science*, vol. 33, pp. 1302–1316, 1987.

[88] S. Axsater, "Modelling emergency lateral transshipments in inventory systems," *Management Science*, vol. 36, pp. 1329–1338, 1990.

[89] M. Dada, "A two-echelon inventory system with priority shipments," *Management Science*, vol. 38, no. 8, pp. 1140–1153, 1992.

[90] C. C. Sherbrooke, "Multiechelon inventory systems with lateral supply," *Naval Research Logistics*, vol. 39, no. 1, pp. 29–40, 1992.

[91] P. Kochel, "On queueing models for some multi-location problems," *International Journal of Production Economics*, vol. 45, no. 1-3, pp. 429–433, 1996.

[92] T. W. Archibald, S. A. E. Sassen, and L. C. Thomas, "An optimal policy for a two depot inventory problem with stock transfer," *Management Science*, vol. 43, no. 2, pp. 173–183, 1997.

[93] P. Alfredsson and J. Verrijdt, "Modeling emergency supply flexibility in a two-echelon inventory system," *Management Science*, vol. 45, no. 10, pp. 1416–1431, 1999.

[94] J. Grahovac and A. Chakravarty, "Sharing and lateral transshipment of inventory in a supply chain with expensive low-demand items," *Management Science*, vol. 47, no. 4, pp. 579–594, 2001.

[95] A. Kukreja, C. P. Schmidt, and D. M. Miller, "Stocking decisions for low-usage items in a multilocation inventory system," *Management Science*, vol. 47, no. 10, pp. 1371–1383, 2001.

[96] H. Wong, D. Cattrysse, and D. V. Oudheusden, "Inventory pooling of repairable spare parts with non-zero lateral transshipment time and delayed lateral transshipments," *European Journal of Operational Research*, vol. 165, no. 1, pp. 207–218, 2005.

[97] H. Wong, G. J. van Houtum, D. Cattrysse, and D. V. Oudheusden, "Multi-item spare parts systems with lateral transshipments and waiting time constraints," *European Journal of Operational Research*, vol. 171, no. 3, pp. 1071–1093, 2006.

[98] H. Zhao, V. Deshpande, and J. K. Ryan, "Emergency transshipment in decentralized dealer networks: When to send and accept transshipment requests," *Naval Research Logistics*, vol. 53, no. 6, pp. 547–567, 2006.

[99] A. A. Kranenburg and G. J. van Houtum, "A new partial pooling structure for spare parts networks," *European Journal of Operational Research*, vol. 199, no. 3, pp. 908–921, 2009.

[100] E. Kutanoglu, "Insights into inventory sharing in service parts logistics systems with time-based service levels," *Computers & Industrial Engineering*, vol. 54, no. 3, pp. 341–358, 2008.

[101] E. Kutanoglu and M. Mahajan, "An inventory sharing and allocation method for a multi-location service parts logistics network with time-based

service levels," *European Journal of Operational Research*, vol. 194, no. 3, pp. 728–742, 2009.

[102] F. V. Utterbeeck, H. Wong, D. V. Oudheusden, and D. Cattrysse, "The effects of resupply flexibility on the design of service parts supply systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 1, pp. 72–85, 2009.

[103] L. Tiacci and S. Saetta, "Reducing the mean supply delay of spare parts using lateral transshipments policies," *International Journal of Production Economics*, vol. 133, pp. 182–191, Sept. 2011.

[104] F. Karsten, M. Slikker, and G.-J. van Houtum, "Inventory pooling games for expensive, low-demand spare parts," *Naval Research Logistics (NRL)*, vol. 59, no. 5, pp. 311–324, 2012.

[105] S. Axsater, "A new decision rule for lateral transshipments in inventory systems," *Management Science*, vol. 49, no. 9, pp. 1168–1179, 2003.

[106] J. Liu and C. Lee, "Evaluation of inventory policies with unidirectional substitutions," *European Journal of Operational Research*, vol. 182, no. 1, pp. 145–163, 2007.

[107] S. Nahmias and W. S. Demmy, "Operating characteristics of an inventory system with rationing," *Management Science*, vol. 27, no. 11, pp. 1236–1245, 1981.

[108] A. Y. Ha, "Stock-rationing policy for a make-to-stock production sys-

tem with two priority classes and backordering," *Naval Research Logistics*, vol. 44, no. 5, pp. 457–472, 1997.

[109] R. Dekker, M. J. Kleijn, and P. J. de Rooij, "A spare parts stocking policy based on equipment criticality," *International Journal of Production Economics*, vol. 56, pp. 69–77, 1998.

[110] F. de Vericourt, F. Karaesmen, and Y. Dallery, "Optimal stock allocation for a capacitated supply system," *Management Science*, vol. 48, no. 11, pp. 1486–1501, 2002.

[111] K. D. Cattani and G. C. Souza, "Inventory rationing and shipment flexibility alternatives for direct market firms," *Production & Operations Management*, vol. 11, no. 4, pp. 441–457, 2002.

[112] V. Deshpande, M. A. Cohen, and K. Donohue, "A threshold inventory rationing policy for service-differentiated demand classes," *Management Science*, vol. 49, no. 6, pp. 683–703, 2003.

[113] M. M. Fadiloglu and O. Bulut, "An embedded markov chain approach to the analysis of inventory systems with backordering under rationing," tech. rep., Working paper, 2005.

[114] O. Vicil and P. Jackson, "An exact optimal solution to a threshold inventory rationing model for multiple priority demand classes," tech. rep., 2006.

[115] H. Arslan, S. C. Graves, and T. A. Roemer, "A single-product inventory model for multiple demand classes," *Management Science*, vol. 53, no. 9, pp. 1486–1500, 2007.

[116] Y. L. Kocaga and A. Sen, “Spare parts inventory management with demand lead times and rationing,” *IIE Transactions*, vol. 39, no. 9, pp. 879–898, 2007.

[117] A. A. Kranenburg and G. J. van Houtum, “Service differentiation in spare parts inventory management,” *Journal of the Operational Research Society*, vol. 59, no. 7, pp. 946–955, 2007.

[118] R. H. Teunter and W. K. K. Haneveld, “Dynamic inventory rationing strategies for inventory systems with two demand classes, poisson demand and backordering,” *European Journal of Operational Research*, vol. 190, no. 1, pp. 156–178, 2008.

[119] P. Enders, I. Adan, A. Scheller-Wolf, and G.-J. v. Houtum, “Inventory rationing for a system with heterogeneous customer classes,” *Flexible Services and Manufacturing Journal*, pp. 1–43, May 2012.

[120] E. Alvarez, M. van der Heijden, and W. Zijm, “The selective use of emergency shipments for service-contract differentiation,” *International Journal of Production Economics*, Mar. 2012.

[121] H. Tiemessen, M. Fleischmann, G. van Houtum, J. van Nunen, and E. Pratsini, “Dynamic demand fulfillment in spare parts networks with multiple customer classes,” *European Journal of Operational Research*, vol. 228, pp. 367–380, Jan. 2013.

[122] E. M. Alvarez, M. C. v. d. Heijden, and W. H. M. Zijm, “Service differentiation in spare parts supply through dedicated stocks,” *Annals of Operations Research*, pp. 1–21, Apr. 2013.

[123] K. Moinzadeh and C. P. Schmidt, "An (S-1,S) inventory system with emergency orders," *Operations Research*, vol. 39, no. 2, pp. 308–321, 1991.

[124] M. van der Heijden, E. Alvarez, and J. Schutten, "Inventory reduction in spare part networks by selective throughput time reduction," *International Journal of Production Economics*, vol. 143, pp. 509–517, Mar. 2012.

[125] E. Ozkan, G.-J. v. Houtum, and Y. Serin, "A new approximate evaluation method for two-echelon inventory systems with emergency shipments," *Annals of Operations Research*, pp. 1–23, June 2013.

[126] S. Axsater, C. Howard, and J. Marklund, "A distribution inventory model with transshipments from a support warehouse," *IIE Transactions*, vol. 45, no. 3, pp. 309–322, 2013.

[127] J. A. Muckstadt, "A model for a multi-item, multi-echelon, multi-indenture inventory system," *Management Science*, vol. 20, no. 4, p. 472, 1973.

[128] A. Svoronos and P. Zipkin, "Evaluation of one-for-one replenishment policies for multiechelon inventory systems," *Management Science*, vol. 37, no. 1, pp. 68–83, 1991.

[129] W. J. Kennedy, J. W. Patterson, and L. D. Fredendall, "An overview of recent literature on spare parts inventories," *International Journal of Production Economics*, vol. 76, no. 2, pp. 201–215, 2002.

[130] V. D. R. Guide and R. Srivastava, "Repairable inventory theory: Models and applications," *European Journal of Operational Research*, vol. 102, no. 1, pp. 1–20, 1997.

[131] G. Bolch, *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications.* Wiley-Blackwell, 2006.

[132] D. R. Cox and W. L. Smith, *Queues.* Chapman & Hall/CRC, 1971.

[133] C. C. Sherbrooke, *Optimal Inventory Modeling of Systems: Multi-Echelon Techniques.* Springer Netherlands, 2004.

[134] S. Stidham, *Optimal Design of Queueing Systems.* CRC Press/Taylor & Francis, 2009.

[135] S. Meyn, *Control Techniques for Complex Networks.* Cambridge University Press, 2008.

[136] A. Mandelbaum and A. L. Stolyar, "Scheduling flexible servers with convex delay costs: heavy-traffic optimality of the generalized c|mu rule," *Operations Research*, vol. 52, no. 6, pp. 836–855, 2004.

[137] J. A. van Mieghem, "Dynamic scheduling with convex delay costs: The generalized c|mu rule," *The Annals of Applied Probability*, vol. 5, pp. 809–833, 1995.

[138] F. De Vericourt, F. Karaesmen, and Y. Dallery, "Dynamic scheduling in a make-to-stock system: A partial characterization of optimal policies," *Operations Research*, vol. 48, no. 5, pp. 811–819, 2000.

[139] A. Perez and P. Zipkin, "Dynamic scheduling rules for a multiproduct make-to-stock queue," *Operations Research*, vol. 45, no. 6, pp. 919–930, 1997.

[140] A. L. Stolyar, "Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic," *Annals of Applied Probability*, pp. 1–53, 2004.

[141] M. Bramson, "State space collapse with application to heavy traffic limits for multiclass queueing networks," *Queueing Systems*, vol. 30, no. 1-2, pp. 89–140, 1998.

[142] H. Chen and A. Mandelbaum, "Hierarchical modeling of stochastic networks, part i: fluid models," in *Stochastic modeling and analysis of manufacturing systems*, pp. 47–105, Springer, 1994.