2013

Proactive and Efficient Spare Parts Inventory Management Policies Considering Reliability Issues

Jingyao Gu
University of Windsor

Proactive and Efficient Spare Parts Inventory Management Policies Considering Reliability Issues

by

Jingyao Gu

A Thesis

Submitted to the Faculty of Graduate Studies

through the Department of Industrial and Manufacturing Systems Engineering

in Partial Fulfillment of the Requirements for

the Degree of Master of Applied Science at the

University of Windsor

Windsor, Ontario, Canada

2013

Proactive and Efficient Spare Parts Inventory Management Policies Considering Reliability Issues

by

Jingyao Gu

APPROVED BY:

____

Dr. Xiaolei Guo
Odette School of Business

____

Dr. Zbigniew Pasek
Department of Industrial and Manufacturing System Engineering

____

Dr. Guoqing Zhang, Co-Advisor
Department of Industrial and Manufacturing System Engineering

____

Dr. Kevin Li, Co-Advisor
Odette School of Business

____

Dr. Reza Lashkari, Chair of Defense
Department of Industrial and Manufacturing System Engineering

March 15, 2013

AUTHOR'S DECLARATION OF ORIGINALITY

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

Spare parts inventory management plays an important role in many industries. They exist to serve the maintenance planning and a good planning can significantly reduce maintenance cost. This thesis developed a series of non-linear programming models to obtain optimal spare parts replenishment policies for failure-based maintenance in a single period. Both single Part Number case and multiple Part Numbers case with a budget constraint are addressed. Compared with traditional forecasting methods which only consider historical data, our proposed inventory policies take into account reliability issues and predict impending demands based on part failure distributions from two perspectives: failure time and failure numbers. Therefore, optimal order quantity and best order time can be found to realize total cost minimization, as well as a systematic inventory optimization.

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my supervisors, Dr. Guoqing Zhang, and Dr. Kevin Li, for their devoted guidance, immense support and constant encouragement during my graduate study in Windsor. Without their excellent mentoring skills and professional knowledge, this work would not have been done. I will eternally remember this life-changing experience to work with them.

I would also like to convey my heartiest appreciation to Dr. Zbigniew Pasek and Dr. Xiaolei Guo for being on my advisory committee and providing valuable suggestions on my work.

Special thanks go to Dr. Guoqing Zhang, Dr. Wagiuh ElMaraghy , Dr. Reza Lashkari, Dr. Chris Lee, and Dr. Saman Hassanzadeh Amin for their generosity in transferring their knowledge which helped me qualify for accomplishing this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 General Overview

Spare parts inventory management plays an important role in many industries, such as airline, trucking, and manufacturing industries. Spare parts are interchangeable parts that are kept in an inventory and used for the repair or replacement of failed parts. The problem of offering an adequate yet efficient supply of spare parts, in support of maintenance and repair of aircraft, trucks, plant and equipment, is an especially vexing inventory management scenario. Most spare parts are very expensive and costly to keep in stock. However, they must be on hand once needed due to high cost of flight cancellation, logistics interruption or plant shutdown.

In recent years, the investments in market of spare parts management are growing. According to Bacchetti and Saccani (2011), the estimation from Aberdeen Research pointed that the total market size for spare parts management software is over 100 millions dollars in 2005. In addition, compared with traditional inventory management, techniques for managing spare parts inventories should be different from the ones used for finished products or components used in production due to the following unique features of spare parts inventories. First of all, work-in-process (WIP) inventories exist to smooth out irregularities in production flows, and finished product inventories are held to deal with fluctuations in customer demand. Whereas the key concern of spare parts inventories is to help keep aircraft, trucks, equipment and plant in an operating condition. Furthermore, WIP and final product inventories can be adjusted by changing production rates, and expediting delivery. However, spare parts inventory levels are mainly determined by how equipment or vehicles are used and how they are maintained. In addition, the spare parts shortage costs are usually very high and sometimes not easy to gauge. Because maintenance parts stockouts may not only lead to significant production losses, but also intangible cost such as increased risk to operating personnel. Finally, obsolescence may also be a problem for spare parts due to their specialized uses.

Extensive research has been conducted in this field. Thomas and Osaki (1978) provided an optimal ordering policy for a spare unit with lead time. Kennedy et al. (2002) analyzed the differences of spare parts inventories and work in process inventories, and discussed the unique aspects of spare parts inventories. Vaughan (2003) proposed a failure replacement and preventive maintenance spare parts ordering policy. Wang (2012) described a stochastic model for joint spare parts inventory and planned maintenance optimization. However, it is still an especially challenging problem because spare parts demands are usually generated by the need of maintenance which is difficult to predict based on historical data of past spare parts usages.

1.2 Proposed Research

1.2.1 Research Topic

Our research was motivated by creating an efficient spare parts inventory model to provide better service for maintenance needs. The objective of this research is to develop an effective approach to determine optimal spare part inventory policies with taking into account reliability issues. In details, this research involves:

1) Analyze the features of spare part inventory management problem with considering the failure and maintenance process. In our research, we take airline industry as an example but the approaches and results proposed are not limited to the industry.
2) Establish the mathematical models to help decision-makers to determine the best order time and the optimal order quantities so that the total cost, including purchasing, holding, and shortage costs, is minimal. The uncertainty on both the number of failures and the lifetime for each part is considered in the problem.
3) Consider both single spare part and multiple spare parts with a budget constraint. It is very common that there are lots of spare parts in an inventory system and the company usually faces budget constraint when making ordering decision for those multiple spare parts. Thus it is more practical to investigate inventory policies for multiple spare parts with a budget constraint.
4) Develop effective solution methods for the proposed model. Since the uncertainty on both the number of failures and the lifetime is considered, the proposed models

are nonlinear programming problems or stochastic programming problems, and they are difficult to solve, especially for multiple Part Number case.

5) Make sensitivity analysis to compare different spare part ordering policies.

1.2.2 Organization of Thesis

This thesis is organized as follows. Chapter 2 describes two single part number (PN) models, a basic mathematical model and an improved mathematical model. The basic model assumes that the shortage period starts from mean time to failure (MTTF). Numerical and iterative methods as well as GAMS are employed to solve this model. The improved model takes into account accurate shortage time. Due to its complexity, only GAMS is applied in solution methodology. Both models are proved effective in cost reduction as reflected by numerical examples and their results. Comparisons of the two models are also discussed.

Chapter 3 explores multiple PNs models with a budget constraint. As the results of GAMS and its solvers for large-scale instances are not reliable, we use a Lagrangian relaxation heuristic to relax the budget constraint and decompose the large sized problem. The advantage of the Lagrangian relaxation heuristic is that it can provide a measure for the gap between the bound of the optimal solution and the approximate solution. To overcome the instability of solvers in GAMS, we further combine an iterative method with the Lagrangian relaxation heuristic, and develop a numerical algorithm to find the optimal objective function value and decision variables. Both small sized and large sized examples show that the gap between the bound and optimal solution is 0. Finally, Chapter 4 provides the conclusions and suggestions for future research.

CHAPTER 2: EFFICIENT SPARE PARTS INVENTORY MANAGEMENT FOR AIRCRAFT MAINTENANCE – SINGLE PART NUMBER (PN) CASE

2.1 Introduction

In airline industries, an operator has to deal with two types of issues: the aircraft operating cost and customer satisfaction. Aircraft maintenance planning plays a major role in both of them. On the one hand, based on an analysis conducted by the International Air Transport Association (IATA)'s Maintenance Cost Task Force, the maintenance cost takes up about 13% of the total operating cost, and it can be reduced by a good planning. On the other hand, an excellent maintenance program can effectively avoid flight delays and cancellations, thus improve customer satisfaction and competitiveness in the industry. Spare parts inventories exist to serve the maintenance planning. An excess of spare parts inventory leads to a high holding cost and impedes cash flows, whereas inadequate spare parts can result in costly flight cancellations or delays with a negative impact on airline performance. Since the airline industry involves with a large number of parts and some of them are quite expensive, it is important to find an appropriate inventory model to achieve a right balance.

Compared with other industries, the airline industry is unique due to a combination of four market characteristics: global need for parts, demand unpredictability, traceability of parts for safety reasons, and high cost of not having a part. Traditionally, spare parts are generally classified into four groups: Rotables, Repairables, Expendables and Consumables. For different categories, different replenishment policies are used. Rotables and Repairables are mainly based on predicted failures estimated by manufacturers, and the planning parameters are finished as management decision. As to Expendables and Consumbles, the reorder point system (ROP) is used and input comes from historical demand with estimated changes. However, this kind of inventory management is typically subjective and imprecise, thus is not an ideal policy. From a survey conducted by Ghobbar and Friend (2004), 152 out of 175 respondents were using the ROP system and about half were dissatisfied and considering implementing new systems.

Our research was motivated by creating an efficient spare parts inventory model in order to provide better service for maintenance needs. When aircraft parts fail, they generate demand for spare parts, and are supplied from spare parts inventory. Under ideal situation, those parts should be in stock and in turn replenished by further activities such as purchasing or repairing. Demands will be satisfied immediately, and aircraft maintenance work can take place on schedule. However, if required spare parts are not available at that time, even purchase orders can be accepted by suppliers at once, delivery time is still a big issue that cannot be ignored. Postponed troubleshooting due to spare parts shortage will probably lead to flight delay or cancellation which will incur huge extra cost. Unfortunately, the second situation is hard to avoid because of uncertain parts failures, large number of parts, limit budget and warehouse space, etc. We try to establish an efficient spare parts inventory model that use minimum expense to achieve maximum productivity. Unlike the previous inventory models that just address the problem of determining the amount of parts to be purchased, our efficient inventory model satisfies spare parts demands from two perspectives: quantity and time. Therefore, it can better improve service level and control the total costs which generally include purchasing cost, holding cost, and shortage cost.

Rotables	Complex components Normally unlimited number of repairs Normally no scrap is expected Controlled by individual serial number Exchange during maintenance
Repairables	components which can be technically and economically repaired: Under normal conditions, a follow up of each individual serial number is not necessary. Have limited number of repairs and also have a possibility of scrap
Expendables	cannot be repaired and will be scrapped after removal and inspection result is unserviceable 100% replacement items Items which cannot be repaired (not economical to be repaired) Standard parts
Consumables	any materials used only once Raw material Chemical material Items which merge on production with new product and cannot be removed

Table 1. Definitions of Rotables, Repairables, Expendables and Consumables

In the context of our model, the installed parts failure distribution is introduced. We assume failures can be predicted based on maintenance data or manufacturer's manual, and maintenance activities are the key drivers of spare parts demand. Advance orders are triggered to reduce downtime caused by parts delivery time. In our analysis, we examine the parts failure distribution to find optimal order time and order quantity by considering that the lifetime and quantity of installed parts failure distribution may influence the duration and numbers of spare parts shortage or overstock, thus result to total cost fluctuation. A non-linear programming (NLP) model is presented with the objective of minimizing air carriers' expected cost in spare parts. Numerical and iteration methods and GAMS are employed to solve the model.

This chapter is organized as follows. In the next section, we give a brief literature review. Section 2.3 presents a basic mathematical model considering shortage period starts from mean time to failure (MTTF). Numerical and iteration methods as well as GAMS can be used to solve this model. We also develop an improved mathematical model, which takes into account exact shortage time, and its solution methodology in Section 2.4. Section 2.5 illustrates the value of our models in cost reduction by numerical examples and their results. Sensitivity analysis and models comparison is conducted in the following section. Finally, section 2.7 provides the conclusions and suggestions for future research.

2.2 Literature Review

Over the past few decades, great efforts have been made to improve inventory management. Among those work, Ghobbar and Friend (2003) discussed the forecasting of intermittent demand in relation to these primary maintenance processes, and compared the experimental results of thirteen forecasting methods. Regattieri et al. (2005) analyzed the behavior of forecasting techniques when dealing with lumpy demand, and made a comparison for twenty forecasting techniques. Both papers found that the best approaches for intermittent demand are weighted moving average, Holt and Croston methods. Furthermore, Campbell (1963) examined demand data from the United State Air Force's maintenance records, and explored relationships between demand and operational variables. He concluded that demand seemed to be related to flying hours and sorties flown, with flying hours having a stronger relationship. Ghobbar and Friend (2002) investigated the source of demand lumpiness, and proposed an assumption that demand is strictly linearly to flying hours/landings. Today, more companies are considering flying hours as the major factor in their forecasting of demand calculation and using the mean time between removal/ overhaul (MTBR/O) to forecast a failure rate. Thus, preventive maintenance (PM) is widely used especially for some critical components that directly affect flight safety. Many papers are presented to address spare parts and failure-based maintenance actions or spare parts with either an age or block-based replacement policy. The earliest papers can be traced to Natarajan (1968) who proposed a reliability problem with spares and Allen and D'esopo (1968) who studied an ordering policy for repairable stock items. Armstrong and Atkins (1996) and de Smidt-Destombes (2007) described the joint optimization of spare parts inventory and age or block-based replacement policies. Kim et al. (1996) addressed a failure-based repair policy and its connection with spare parts provision, focusing on how equipment failures affect the spare parts inventory policy without considering the influence of PM. Vaughan (2003) proposed a failure replacement and PM spare parts ordering policy.

The most recent paper in this subject area is Wang (2012), who presents a model to optimize the order quantity, order intervals, and PM intervals jointly under a two-stage failure process. The aforesaid papers mainly address the problem either from an inventory point of view based on the past spare parts usages to forecast the future demand, or from a PM point of view to find an optimal order quantity and PM interval. In practice, spare parts demands are highly related with flying hours/landings. Due to the correlation between part aging and failures, impending high demands for a part might be forecasted even the current demand is low, which is counter to the traditional replenishment system that high demand triggers replenishment and low demand scales back replenishment. This partially explains why the forecasting methods based on historical data cannot be directly used for forecasting future demand in our case. Besides, PM inventory management is different from failure-based inventory management. To the author's best knowledge, limited research handles failure-based procurement inventory management which is very common in practice. As the spare parts demand is uncertain, and sometimes the part delivery time may be very long, it could lead significant loss if a critical part fails but there is no spare to replace it.

Deshpande et al. (2006) explored this issue. To improve the performance of aircraft service parts supply chain in the United States Coast Guard (USCG), they used mathematical programming tools to link the demand transactions to a corresponding maintenance activity. Subsequently, they developed an approach to use part-age data to make inventory decisions. It sets an age threshold and observes the number of installed parts whose age is greater than the threshold, thereby deciding the advance order quantity. This approach tries to synchronize the inventory of good parts with demand distributions, and replenish the inventory just as anticipated demands arrive. It has great advantages compared with traditional inventory policies. However, one important operational problem is not mentioned- when is the best time to issue orders? Ordering at the beginning of period will result in high holding cost, whereas replenishing at the end of period may lead to extensive shortage cost, both tend to drive up the total cost. Our efficient inventory model considers both order time and order quantity. Furthermore, in Deshpande et al. (2006), based on the assumption that lead-time demand D and the signal level S follow a joint bivariate-normal distribution, they derived a result that the total cost per unit time is minimized by setting the part-age threshold T to a value that maximizes the correlation $\rho(T)$ between D and S. In contrast, we introduce the parts failure distribution from two aspects, life time and total number of failures, and assume they are uncorrelated, that is, ρ equals to 0. Because all demands come from installed parts failures, we can predict impending demands and develop an efficient proactive inventory model to replenish spare parts inventory before most failures occur. Accordingly, the best order time and the optimal order quantities can be worked out by minimizing the total cost which consists of purchasing, holding, and shortage costs.

2.3 Basic Mathematical Model and Solution Methodology

2.3.1 Basic Mathematical Model

In this chapter, we consider a generalized ordering policy for only one kind of part with a given part number (PN) in a single period. A part number is a fundamental identifier of a particular part design used in the airline industry. It unambiguously identifies a part design within a single corporation, or sometimes across several corporations. For example, when specifying a bolt, it is easier to refer to "PN BACB30LH3K24" than describing the key information of the bolt, such as dimensions, material, installed position and manufacturer, which may be lengthy and incomplete. Moreover, multiple parts with the same PN are often found in one or more aircrafts. For instance, if one Boeing 737-300 has installed 200 PN BACB30LH3K24 bolts, and the fleet size of Boeing 737-300 is 20, thus the total number of PN BACB30LH3K24 operated by the carrier will be 4000.

The length of the planning horizon is denoted by T ($0 \leq T \leq \infty$) and the order quantity in this period is denoted by Q. The spare parts for replacement can be delivered after a constant lead time L. The demand is uncertain, and depends on the parts failure distribution. We assume that the number of failures in period T follows a probability density function (PDF) g (·) and a cumulative distribution function (CDF) G (·). The lifetimes of the operating parts are assumed independent with an identical probability density function f (·) and a cumulative distribution function F (·). We also assume g (·) and f (·) are uncorrelated.

Input parameters and function:

h the unit holding cost per unit time

s the unit shortage cost per unit time

T planning horizon, can be infinite

L order lead time

z demand quantity, a random variable

c unit cost

$f(x)$ the PDF of failure distribution considering lifetime for each part

$g(z)$ the PDF of failure distribution considering number of failures for each PN

$F(x)$ the CDF of failure distribution considering lifetime for each part

$G(z)$ the CDF of failure distribution considering number of failures for each PN

Define the following decision variables

t_1 point in time to place an order

t_2 the parts arrival time

Q order quantity

The objective of minimizing the expected total cost is formulated as:

$$Min\ R = h(T-t_2)\int_0^Q (Q-z)g(z)dz + s\left[T-\int_0^\infty xf(x)dx\right]\int_Q^\infty (z-Q)g(z)dz + \left[h\int_{t_2}^\infty (x-t_2)f(x)dx + s\int_0^{t_2}(t_2-x)f(x)dx\right]Q + cQ \tag{1}$$

In the whole planning horizon, we just order this given PN once. If the order quantity is above the total actual demand level, that is, the total number of failures in the whole planning horizon is less than the order quantity Q, the holding cost of those extra stock will start from t_2 and last till the end of the period. The expected holding cost of this part is represented by the first term. Conversely, if the order quantity Q is below the total number of failures, the parts shortage situation will last till the end of the planning horizon. The second term describes this expected shortage cost. Notice that the duration of parts shortage is decided by when the Q^{th} failure occurs and when the planning horizon finishes. This basic model simplifies the problem by using the mean time to failure (MTTF), which is defined by MTTF=$\int_0^\infty xf(x)dx$, to replace the Q^{th} failure time. The third term depicts the expected value of the remaining holding cost and shortage cost during the planning horizon when the order quantity "Q" just matches the total number of failures. Figure 2 illustrates how this part of holding cost and shortage cost are generated. For example, if five failures occurred during the whole planning horizon T, and purchased parts arrived between the third failure and the fourth failure, shortage cost would be incurred due to the first three failures. On the other hand, the remaining ordered parts would be kept in stock and continuously generate holding cost until they are used up. The last term accounts for the purchasing cost. Once we find the optimal parts arrival time, the optimal timing to place an order can be calculated easily by $t_1 = t_2 - L$.

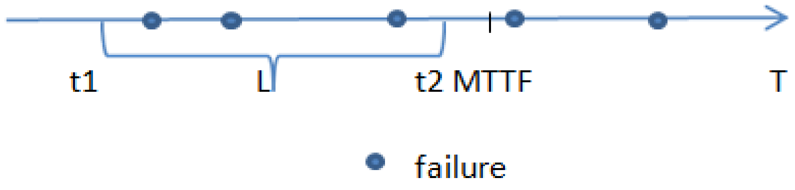

Figure 1. Time sequence

2.3.2 Solution Methodology

2.3.2.1 Numerical and iteration methods

From objective function (1), we are obviously interested in determining the value of t_2 and Q, which minimize the expected cost R. Without considering any constraint, Q and t_2 can be found by the following procedure:

$$\frac{\partial R}{\partial Q} = h(T-t_2)G(Q) - s\left[T - \int_0^\infty xf(x)dx\right][1-G(Q)] + \left[h\int_{t_2}^\infty (x-t_2)f(x)dx + s\int_0^{t_2}(t_2-x)f(x)dx\right] + c \quad (2)$$

It follows that

$\frac{\partial^2 R}{\partial Q^2} = \{h(T-t_2) + s[T - \int_0^\infty xf(x)dx]\}\, g(Q) \geq 0$, for all Q$\geq 0$.

Because the second-order derivative is nonnegative, the function R (Q) is said to be convex. The optimal solution, Q*, occurs where $\frac{\partial R}{\partial Q}$ equals zero. That is,

$$G(Q^*) = \frac{s[T - \int_0^\infty xf(x)dx] - \left[h\int_{t_2}^\infty (x-t_2)f(x)dx + s\int_0^{t_2}(t_2-x)f(x)dx\right] - c}{h(T-t_2) + s[T - \int_0^\infty xf(x)dx]}$$

(3)

Also,

$$\frac{\partial R}{\partial t_2} = -h\int_0^Q (Q-z)g(z)dz - Qh[1-F(t_2)] + QsF(t_2)$$

(4)

It follows that

$\frac{\partial^2 \mathrm{R}}{\partial {t_2}^2} = \ Qf(t_2)(h+s) \geq 0$, for all Q$\geq$ 0.

As the second-order derivative is nonnegative, R (t_2) is convex, and the optimal solution, t_2*, is attained when $\frac{\partial R}{\partial t_2}$ equals zero. That is,

$$F({t_2}^*) = \frac{h\int_0^Q (Q-z)g(z)dz + Qh}{Q(h+s)}$$

(5)

One of the widely-used probability distributions in reliability to model fatigue and wear-out phenomena is the normal distribution, as illustrated by the works of Deshpande et al. (2006), Muchiri and Smit (2011), Tuomas et al. (2001), Batchoun et al. (2002), Byington et al. (2002), and Kiyak (2012). If we assume that f(x) is normally distributed, with a mean μ_x and standard deviation σ_x, the formula for the PDF is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x}\exp\left[-\frac{1}{2}\frac{(x-\mu_x)^2}{{\sigma_x}^2}\right], \quad -\infty < x < \infty$$

Meanwhile, assume g(z) follows a normal distribution with a mean μ_z and standard deviation σ_z. The formula for the PDF is

$$g(z) = \frac{1}{\sqrt{2\pi}\sigma_z}\exp\left[-\frac{1}{2}\frac{(z-\mu_z)^2}{{\sigma_z}^2}\right], \quad -\infty < z < \infty$$

As the random variable ranges from $-\infty$ to $+\infty$, the normal distribution is not a true reliability distribution. However, if for most observed values of mean and standard deviation in the context of this study, the probability that the random variable takes on negative values is negligible, then the normal distribution can be regarded as a reasonable approximation to a failure process. We assume $0 < x < \infty, 0 < z < \infty$, also f(x) and g(z) are uncorrelated.

From equation (3), we have

$$Q^* = G^{-1}\left\{\frac{s(T-\mu_X)-h(\mu_X-t_2)-(h+s)\left[(t_2-\mu_X)F_s\left(\frac{t_2-\mu_X}{\sigma_X}\right)+\sigma_X f_s\left(\frac{t_2-\mu_X}{\sigma_X}\right)\right]-c}{h(T-t_2)+s(T-\mu_X)}\right\} \tag{6}$$

Here $f_s(\cdot)$ is the standard normal density function, and $F_s(\cdot)$ is the standard normal cumulative distribution function for the normal density function $f(x)$. $G^{-1}(\cdot)$ is the inverse of the cumulative density function for the normal density function $g(z)$.

From equation (5), we have

$$t_2{}^* = F^{-1}\left\{\frac{h(Q-\mu_Z)G_s\left[\frac{(Q-\mu_Z)}{\sigma_Z}\right]+h\sigma_Z g_s\left[\frac{(Q-\mu_Z)}{\sigma_Z}\right]+Qh}{Q(h+s)}\right\} \tag{7}$$

Here $g_s(\cdot)$ is the standard normal density function, and $G_s(\cdot)$ is the standard normal cumulative distribution function for the normal density function $g(z)$. $F^{-1}(\cdot)$ is the inverse of the cumulative density function for the normal density function $f(x)$.

Because Q* and $t_2{}^*$ cannot be determined in closed forms from (6) and (7), a numerical algorithm is employed to find solutions. The algorithm converges in a finite number of iterations, provided that a feasible solution exists. The algorithm is described as follows:

Step 0. Use the initial solution

$Q_1 = Q^* = G^{-1}\left\{\frac{s(T-\mu_x)-h(\mu_x-t_2)-(h+s)\left[(t_2-\mu_x)F_s\left(\frac{t_2-\mu_x}{\sigma_x}\right)+\sigma_x f_s\left(\frac{t_2-\mu_x}{\sigma_x}\right)\right]-c}{h(T-t_2)+s(T-\mu_x)}\right\}$, and let $t_2^{(0)} = 0$. Set i=1, and go to step i.

Step i. Use Q_i to determine $t_2^{(i)}$ from equation (7). If $t_2^{(i)} \approx t_2^{(i-1)}$, stop; the optimal solution is $Q^* = Q_i$, and $t_2^* = t_2^{(i)}$. Otherwise, use $t_2^{(i)}$ in equation (6) to compute Q_i. Set i=i+1, and repeat step i.

When the iteration terminates, we can find the optimal timing to place an order as $t_1 = {t_2}^* - L$.

From equation (1), the objective function can be reformulated as:

$$Min\ R = h(T - t_2)\left\{(Q-\mu_z)G_s\left[\frac{(Q-\mu_z)}{\sigma_z}\right] + \sigma_z g_s\left[\frac{(Q-\mu_z)}{\sigma_z}\right]\right\}$$

$$+s(T-\mu_x)\left\{(\mu_z - Q)\left[1 - G_s(\frac{Q-\mu_z}{\sigma_z})\right] + \sigma_z g_s\left[\frac{(Q-\mu_z)}{\sigma_z}\right]\right\}$$

$$+h(\mu_x - t_2)Q + (h+s)Q\left\{(t_2-\mu_x)F_s\left[\frac{(t_2-\mu_x)}{\sigma_x}\right] + \sigma_x f_s\left[\frac{(t_2-\mu_x)}{\sigma_x}\right]\right\} + cQ \tag{8}$$

2.3.2.2 Solve the basic model by GAMS

Another method is to use the function errorf(·) in GAMS to implement the non-linear integral component of the objective function and find an optimal solution. Both MINOS and

CONOPT can yield good results. MINOS is suitable for large constrained problems with a linear or nonlinear objective function and a mixture of linear and nonlinear constraints. For nonlinear constraints, MINOS implements a sequential linearly constrained algorithm derived from the Robinson's method. CONOPT is a feasible path solver based on the generalized reduced gradient method and is often preferable for nonlinear models where feasibility is difficult to achieve.

2.4 An Improved Mathematical Model and Solution Methodology

In the basic model presented in the previous section, the value of T minus MTTF instead of the Q^{th} failure time is adopted to define the parts shortage period till the end of the planning horizon. The improved mathematical model herein aims to find when the Q^{th} failure occurs and plugs it into the model. Therefore, this improved model is designed to find more accurate order quantity Q and order time t_2.

In reliability engineering, it is well known that given that $\mathrm{t}_1, \mathrm{t}_2, \dots, \mathrm{t}_n$, where $\mathrm{t}_\mathrm{i} \le \mathrm{t}_{\mathrm{i}+1}$, are n ordered failure times comprised in a random sample, the number of units surviving at time t_i is n-i. A possible estimate for the reliability function can be expressed as

$$\hat{R}(\mathrm{t}_\mathrm{i}) = \frac{n-i}{n} = 1 - \frac{i}{n}.$$

The estimate for the cumulative failure distribution is

$$\hat{F}(\mathrm{t}_\mathrm{i}) = 1 - \hat{R}(\mathrm{t}_\mathrm{i}) = \frac{i}{n}.$$

If we assume the total number of parts with a given PN in the observed fleet is n, which is technical information provided by the original equipment manufacturer (OEM), and t_Q is the Q^{th} failure time. We also assume that f(x) is normally distributed, and the failure times follow $t_i \sim N(\mu_x, {\sigma_x}^2)$, then

$$F(t_i) = \Phi(\frac{t_i - \mu_x}{\sigma_x}).$$

Because $\hat{F}\left(\mathrm{t}_Q\right) = \frac{Q}{n}$, and $F\left(\mathrm{t}_Q\right) = \Phi(\frac{\mathrm{t}_Q - \mu_x}{\sigma_x})$,

we have

$\frac{Q}{n} = \Phi(\frac{\mathrm{t}_Q - \mu_x}{\sigma_x})$. That is, $Q = n\,\Phi\left(\frac{\mathrm{t}_Q - \mu_x}{\sigma_x}\right)$.

Accordingly, the Q^{th} failure time can be expressed as

$$\mathrm{t}_Q = \sigma_x * F_S^{-1}\left(\frac{Q}{n}\right) + \mu_x, \tag{9}$$

and equation (1) can be improved as

$$\textit{Min } R = h(T - t_2)\int_0^Q (Q - z)g(z)dz + s[T - \sigma_x * F_S^{-1}\left(\frac{Q}{n}\right) - \mu_x]\int_Q^{\infty}(z - Q)g(z)dz + \left[h\int_{t_2}^{\infty}(x - t_2)f(x)dx + s\int_0^{t_2}(t_2 - x)f(x)dx\right]Q + cQ. \tag{10}$$

Compared with Equation (1), Equation (10) is much more complex due to the inverse function $F_S^{-1}\left(\frac{Q}{n}\right)$. Therefore the determination of Q* and ${t_2}^*$ by numerical and iteration methods is not easy. We use GAMS and its solver CONOPT to solve the new objective function (10). Numerical examples and comparative results of the two models are furnished in Section 2.5.

2.5 Numerical Examples and Results

A numerical example, which is introduced in Deshpande et al. (2006), is modified by introducing the distribution of the number of failures for a specified PN. The data are originally drawn from the aircraft maintenance and inventory databases of the United States Coast Guard (USCG) (Deshpande et al., 2006). Here we list the same parameter values used in Deshpande et al. (2006) in Table 2, which is about the main gearbox of aircraft type HH65A.

Note that if we assume a daily flying time of 10 hours, the gearbox mean age at failure should be 2436/10=243.6 days, similarly the standard deviation should be 659/10=65.9 days. Additional parameter values in Table 3 are introduced for our new models.

Unit price	Unit holding cost per day	Unit shortage cost per day	Mean time to failure	Standard deviation
c	h= 0.25*c/365	s=5*c/365	μ_x	σ_x
$449,586.00	$307.94	$6,158.71	2436 hours	659 hours

Table 2. Parameter values of main gearbox in Deshpande et al. (2006)

Planning horizon	Mean number of failures	Standard deviation	Total number of parts observed in fleet
T	μ_z	σ_z	n
5 years=1,825days	25	10	200

Table 3. New designed parameter values for main gearbox

Table 4 summarizes the calculation results of both the Iterative method and GAMS for the basic model, and of GAMS for the improved model. We can see that, for the basic model, the decision variables for both the Iterative approach and the GAMS approach produce very similar results. The objective function values for both approaches are slightly different with a percentage error margin of 0.45%. This error is likely due to the assumption that $\int_{x=0}^{a} xf(x)dx \cong \int_{x=-\infty}^{a} xf(x)dx = \mu F_s\left[\frac{(a-\mu)}{\sigma}\right] - \sigma f_s\left[\frac{(a-\mu)}{\sigma}\right]$ in the iterative approach, which neglects the part of the negative values in the normal distribution. On the other hand, both approaches yield almost identical decision variable values. Next we compare the differences of GAMS results between the basic and improved models. The values of t_2 are essentially the same for the models. However, the values of Q and R change to a certain degree. Compared with the basic model, the value of Q for the improved model increases by 0.55%, and the value of R increases by 0.54%. A closer examination of the two objective functions reveals that the only difference exists in the second term: the shortage period described in the basic model is $(T - \mu_x)$, while in the improved model it is $(T - t_Q) = \left[T - \sigma_x * F_S^{-1}\left(\frac{Q}{n}\right) - \mu_x\right]$. Because the change is only related to Q and R, it barely affects the optimal value of t_2. In this example, t_Q is less than μ_x implying that the shortage situation starts a little earlier in the improved model than that in the basic model, therefore more shortage cost would be incurred and more spare parts should be ordered during the planning horizon.

In order to gain better insights into the proposed models and understand their values for cost reduction in reality, we input different values of Q and t_2 in equation (8) and then compare their objective function values. The basic model is chosen because its format is easier than the improved model whereas the calculation results of both models are comparable as illustrated in Table 4. Figure 2 illustrates the relationship between the expected cost and parts arrival time for the main gear box when Q=37.90. The optimal value of parts arrival time, which minimizes the expected cost, should be set at 143.52. Based on the trend, as the parts arrival time increases, the expected cost first decreases slightly, followed by a dramatic increase. The reason is that, compared with the shortage cost, the holding cost only takes a small fraction of the unit cost. Moreover, the mean age at failure happens at the early period of the planning horizon.

Solution approach		**Iterative**	**GAMS**	
		basic model	**basic model**	**improved model**
Objective function value	R	$30,110,394.24	$29,974,161.85	$30,135,359.75
Decision variable	Q	37.90	37.92	38.13
	t_2	143.52	143.41	143.48
	t_Q	*	*	185.91
Iteration number		3	14	14
Feasible solution		yes	yes	yes

Table 4. Results from iterative and GAMS solution approaches

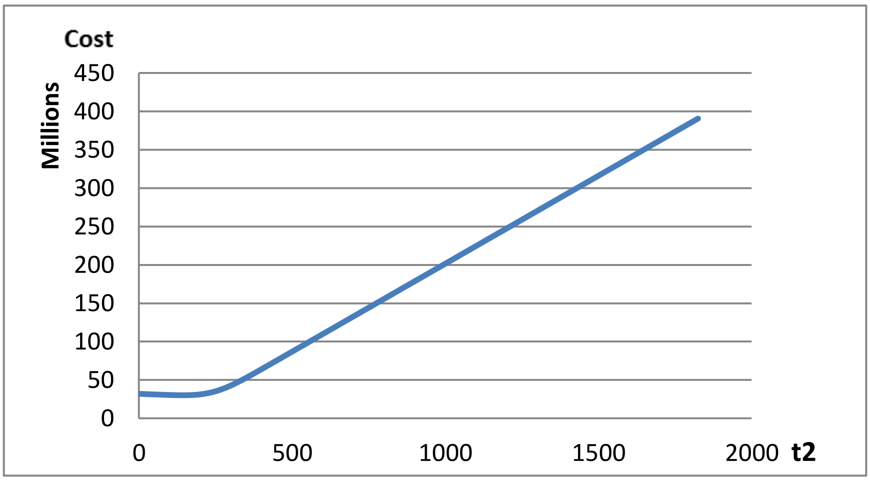

Figure 2. Cost vs. parts arrival time for main gearbox when Q=37.90

Figure 3 depicts the relationship between the expected cost and order quantity curve for the main gear box when t_2=143.52. From the figure, we can find that as the order quantity increases, the expected cost drops sharply till Q= 37.90, then followed by a much slower gradual increase. The different slopes of the curve can be intuitively explained as follows: if the actual order quantity is below the optimal order quantity, compared with overstocking, the shortage cost is much higher than the holding cost.

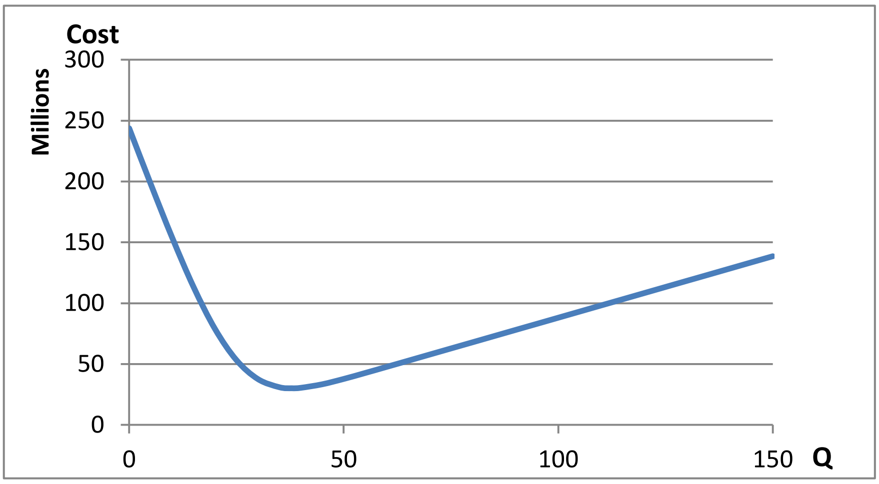

Figure 3. Cost vs. order quantity for gearbox when t_2=143.52

Decision variable		Objective function value	% Cost reduction
t_2	Q	Total cost R	(R-Optimal value)/R
143.515	37.90	$30,110,394.24	Optimal value
0	0	$243,691,145.51	87.64%
143.515	0	$243,690,259.81	87.64%
143.515	150	$138,579,282.00	78.27%
0	37.90	$31,923,057.35	5.68%
1825	37.90	$390,709,822.09	92.29%
1825	150	$1,528,346,051.51	98.03%

Table 5. Representative points in Figure 1 and Figure 2

We also select some representative points in Figures 2 and 3, and list them in Table 5. By comparing the total cost of each point with the optimal value, we can see that the optimal policy leads to a significant reduction in the total inventory cost, ranging from 5.68% to 98.03%. It is further noted that if the order quantity and arrival time deviate from the optimal solution, early arrival is preferred to late arrival due to low holding cost, high shortage cost, and early failures.

To further explore how the order time affects the inventory replenishment policy, we solve another example by modifying h, s, and μ_x. The updated parameter values are listed in Table 6, remaining parameters assume the same values as those in the first example. We summarize the calculation results in Table 7.

Unit holding cost per day	Unit shortage cost per day	Mean time to failure
h= 0.5*c/365	s=2*c/365	μ_x
$615.87	$2,463.48	12180 hours

Table 6. Modified parameter values

Solution approach		Iterative	GAMS	
		basic model	basic model	improved model
Objective function value	R	$20,234,054.82	$20,161,979.65	$20,787,748.91
Decision variable	Q	25.52	25.55	26.74
	t_2	1,170.03	1,169.75	1170.48
	t_Q	*	*	1144.92
Iteration number		4	15	28
Feasible solution		yes	yes	yes

Table 7. Calculation results for modified parameter values

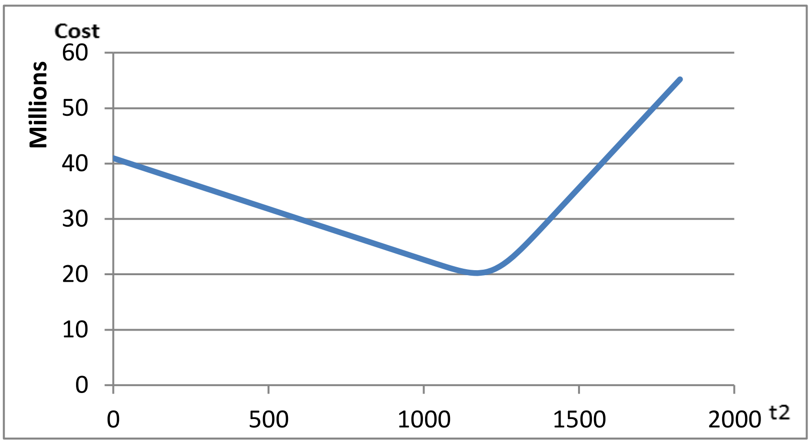

Figure 4. Cost vs. parts arrival time for main gearbox when Q=25.52

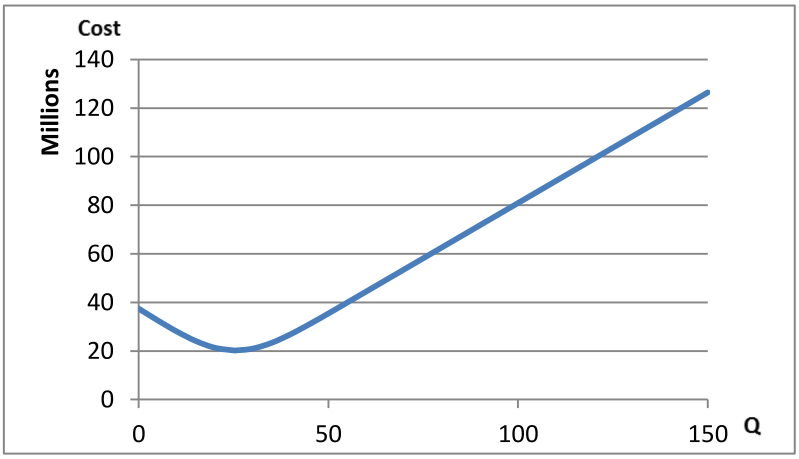

Figure 5. Cost vs. order quantity for gearbox when t_2=1170.03

Decision variable		Objective function value	% Cost reduction
t_2	Q	Total cost R	(R-Optimal value)/R
1170.03	25.52	$20,234,054.82	Optimal value
0	0	$37,435,878.21	45.95%
1170.03	0	$37,421,436.61	45.93%
1170.03	150	$126,438,628.76	84.00%
0	25.52	$40,981,710.29	50.63%
1825	25.52	$55,216,737.88	63.36%
1825	150	$291,738,203.01	93.06%

Table 8. Representative points in Figure 4 and Figure 5

Figure 4 demonstrates the relationship between expected cost and parts arrival time when Q=25.52. Figure 5 shows the expected cost and the order quantity curve when t_2=1170.03. Table 8 provides the comparison results of representative points in Figure 4 and Figure 5. Based on these results, we can find that the optimal policy can contribute to a cost reduction ranging from 45.93% to 93.06%. When the unit holding cost per day (h) doubles, the unit shortage cost per day (s) decreases from 5*c/365 to 2*c/365, and the MTTF (μ_x) lasts 5 times longer than before, even the order quantity is optimal, placing order at the optimal time would save 50.63% cost compared with ordering at the beginning, and 63.36% compared with ordering at the end of the planning horizon. Next, we shall further conduct sensitivity analyses of the basic and improved models by changing certain parameter values.

2.6 Sensitivity and Comparative Analyses

To examine how critical parameters in the model affect the optimal solution, we conduct sensitivity analyses for five different cases and compare their final results between the basic model and improved model. The parameter values are based on the second example in Section 2.5. In each case, only one parameter is changed while the others are kept constant. We investigate how the optimal solutions (total cost R, order quantity Q, parts arrival time t_2) are affected by the failure distributions (μ_x, σ_x, μ_z, σ_z), and how R, Q, and t_2 in the improved model change with the total number of observed parts (n).

First, we calculate the optimal objective values by varying μ_x , the MTTF of given parts. Figure 6 shows that the expected total cost R of both the basic and improved models decreases as μ_x increases from 200 to 1200. Also, the values of R in the improved model are always higher than those in the basic model, at a small margin from 1.04% to 2.90%. Figure 7 illustrates that the optimal order quantity Q in both models decreases when μ_x increases. Compared with the basic model, the values of Q in the improved model are higher, at a margin from 1.05% to 4.24%. Figure 8 shows that the parts arrival time t_2 in both models increases when μ_x increases. The values of both t_2 are almost identical, verifying that the assertion in Section 2.5 that the two models mainly differ in their handling of Q and R, and, hence the optimal value of t_2 is barely affected.

	Basic model			**Improved model**			**Percentage Error**		
μ_x	R	Q	t_2	R	Q	t_2	R	Q	t_2
200	30080101	30.13	154.83	30395999	30.45	155.03	1.04%	1.05%	0.13%
400	28282093	29.72	354.36	28620638	30.09	354.60	1.18%	1.24%	0.07%
600	26433376	29.18	554.01	26802664	29.63	554.30	1.38%	1.52%	0.05%
800	24530299	28.45	753.55	24944128	29.02	753.91	1.66%	1.96%	0.05%
1000	22536336	27.40	952.89	23020242	28.17	953.37	2.10%	2.71%	0.05%
1200	20370771	25.75	1151.87	20980234	26.89	1152.57	2.90%	4.24%	0.06%

Table 9. Objective values comparison of two models when μ_x changes

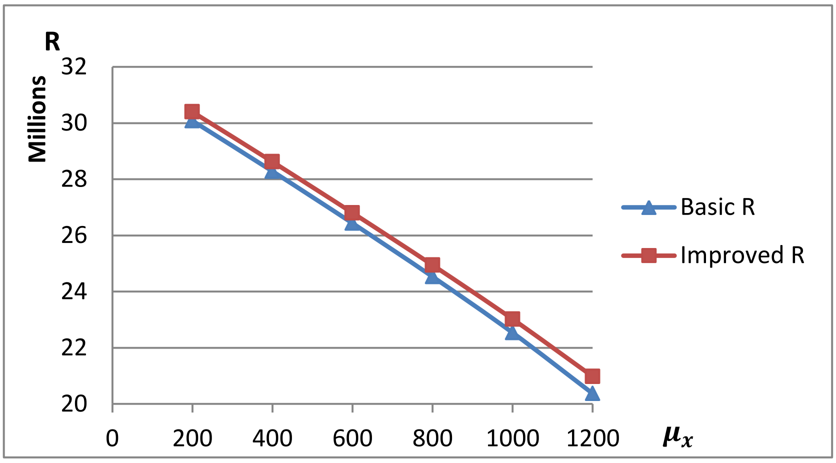

Figure 6. Total cost as μ_x increases

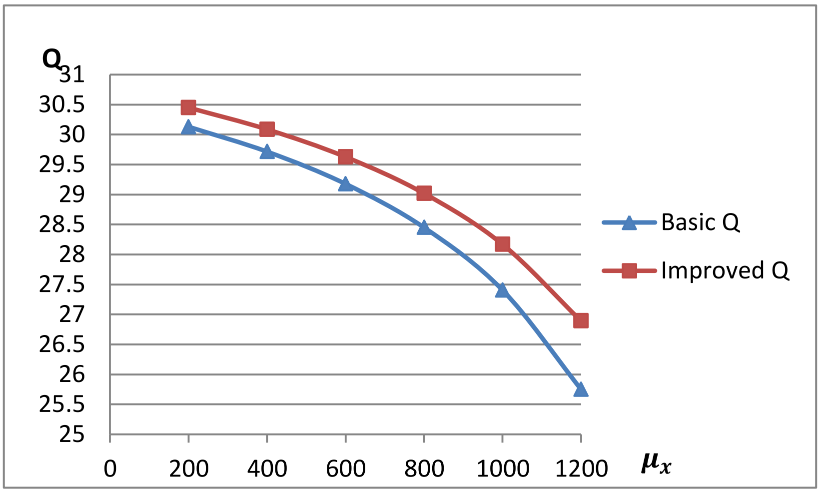

Figure 7. Optimal order quantity as μ_x increases

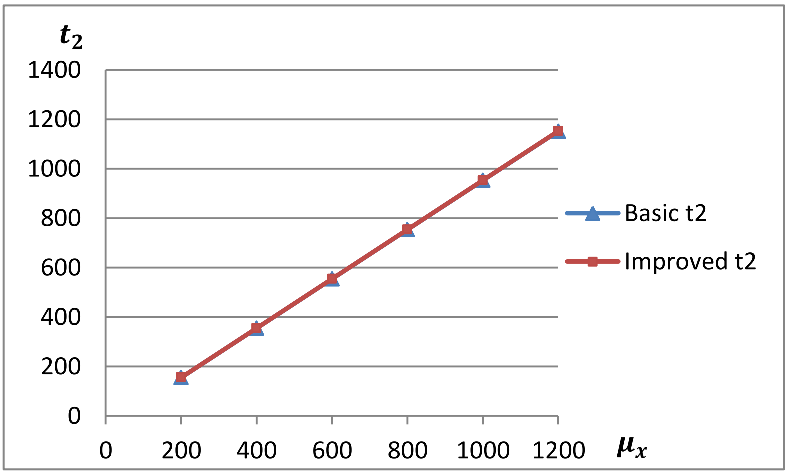

Figure 8. Optimal inventory replenishment time as μ_x increases

	Basic model			Improved model			Percentage Error		
σ_x	R	Q	t_2	R	Q	t_2	R	Q	t_2
20	19042951	26.23	1203.48	19225695	26.60	1203.55	0.95%	1.40%	0.01%
40	19534770	25.93	1188.85	19906533	26.67	1189.13	1.87%	2.76%	0.02%
60	20020055	25.64	1174.12	20587051	26.73	1174.72	2.75%	4.08%	0.05%
80	20498901	25.35	1159.28	21267288	26.78	1160.34	3.61%	5.36%	0.09%
100	20971401	25.06	1144.34	21947276	26.83	1145.97	4.45%	6.62%	0.14%
120	21437640	24.77	1129.3	22627043	26.88	1131.62	5.26%	7.84%	0.21%
140	21897695	24.49	1114.16	23306612	26.92	1117.28	6.05%	9.04%	0.28%
160	22351642	24.21	1098.93	23986003	26.97	1102.95	6.81%	10.23%	0.36%
180	22799548	23.93	1083.61	24665234	27.00	1088.64	7.56%	11.38%	0.46%
200	23241476	23.65	1068.2	25344320	27.04	1074.33	8.30%	12.54%	0.57%

Table 10. Objective values comparison of two models when σ_x changes

Next, we consider the impact on the optimal objective value by changing the value of standard deviation (σ_x) of the parts lifetime. Figure 9 illustrates that the expected total cost R of both the basic and improved models increases as σ_x increases from 20 to 200. This result is natural as a heightened uncertainty level tends to result in more holding and shortage costs. Figure 10 describes an interesting situation that the optimal order quantity Q in the basic model decreases whereas that in the improved model increases when σ_x grows. The order quantity Q in the basic model is always higher than in the improved model. The reason might be that failures between t_Q and μ_x are ignored in the basic model, leading to a lower order quantity than that in the improved model. Furthermore, when σ_x increases, the failures distribution becomes flatter, accordingly more parts will be ordered earlier to guarantee the same service level. However, the optimal value of Q in the basic model decreases due to a growing number of neglected failures between t_Q and μ_x. Finally, Figure 11 confirms that t_2 still has little change.

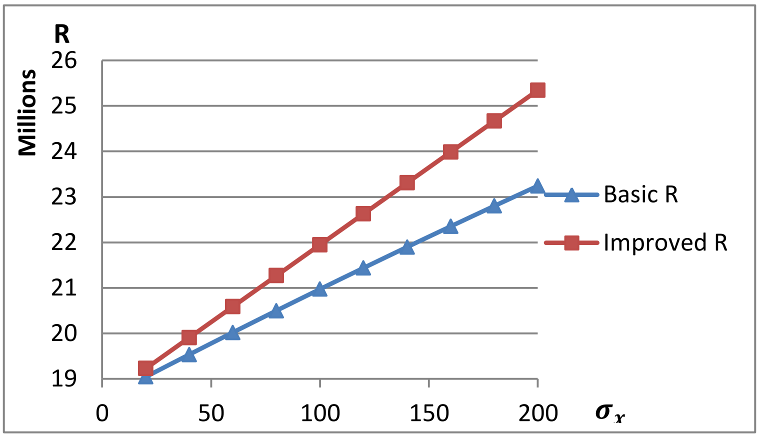

Figure 9. Total cost as σ_x increases

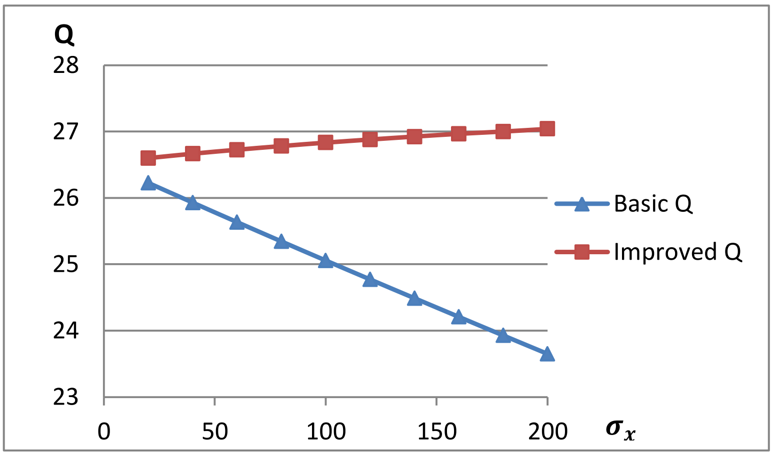

Figure 10. Optimal order quantity as σ_x increases

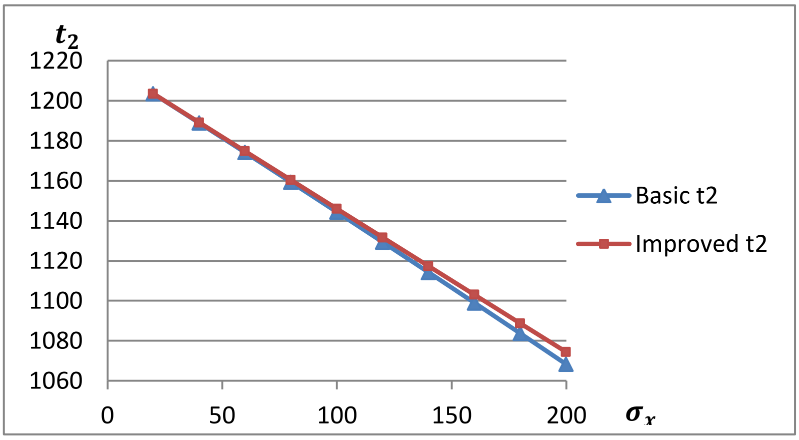

Figure 11. Optimal inventory replenishment time as σ_x increases

When we increase μ_z from 20 to 120, the percentage errors of R, Q , and t_2 between the two models decrease.

	Basic model			Improved model			Percentage Error		
μ_z	R	Q	t_2	R	Q	t_2	R	Q	t_2
20	17476989	20.64	1170.70	18157301	21.95	1171.58	3.75%	5.94%	0.07%
40	27833081	40.51	1167.33	28304380	41.42	1167.75	1.67%	2.19%	0.04%
60	37963663	60.50	1165.78	38263091	61.11	1165.98	0.78%	0.99%	0.02%
80	48092679	80.50	1164.98	48238099	80.84	1165.07	0.30%	0.42%	0.01%
100	58221282	100.50	1164.50	58217152	100.59	1164.52	-0.01%	0.09%	0.00%
120	68349675	120.50	1164.18	68190701	120.34	1164.15	-0.23%	-0.14%	0.00%

Table 11. Objective values comparison of two models when μ_z changes

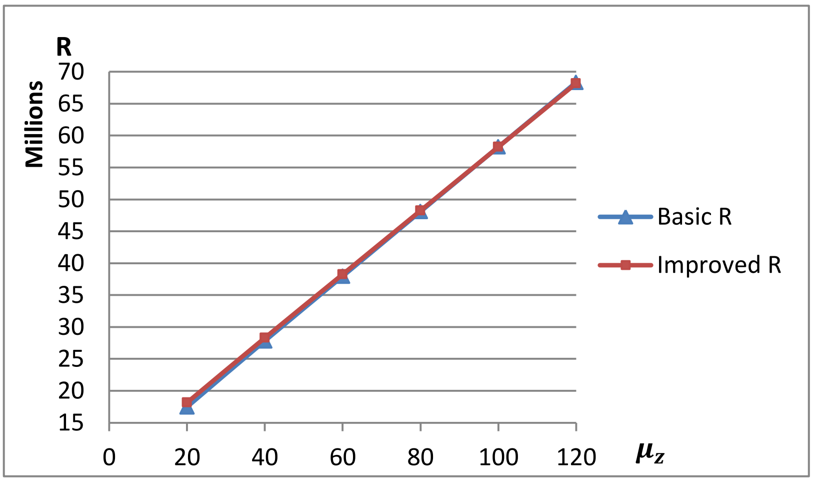

Figure 12. Total cost as μ_z increases

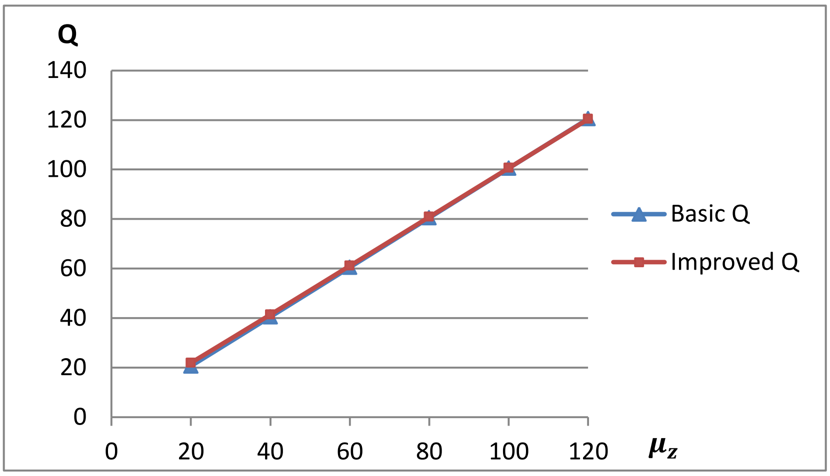

Figure 13. Optimal order quantity as μ_z increases

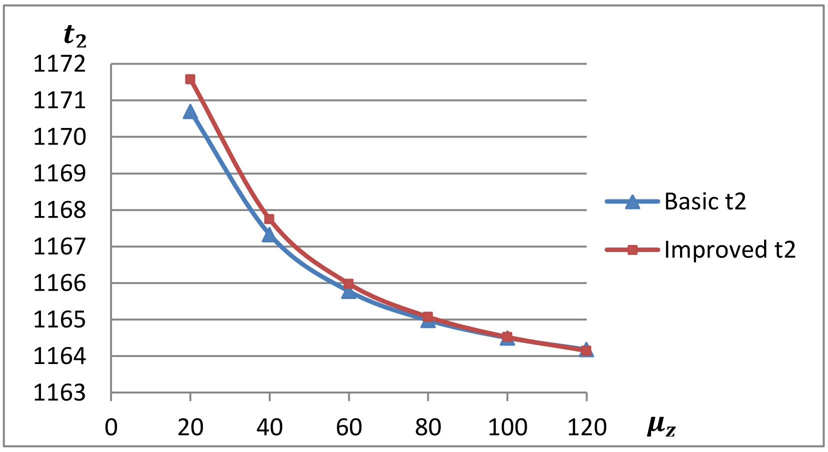

Figure 14. Optimal inventory replenishment time as μ_z increases

	Basic model			Improved model			Percentage Error		
σ_z	R	Q	t_2	R	Q	t_2	R	Q	t_2
2	14176351	25.10	1164.11	14305746	25.32	1164.32	0.90%	0.87%	0.02%
4	15691930	25.20	1165.65	15948873	25.65	1166.03	1.61%	1.77%	0.03%
6	17206557	25.31	1167.15	17589213	26.00	1167.67	2.18%	2.66%	0.04%
8	18710829	25.41	1168.57	19217067	26.35	1169.20	2.63%	3.57%	0.05%
10	20161980	25.55	1169.75	20787749	26.74	1170.48	3.01%	4.46%	0.06%
12	21521340	25.74	1170.56	22259759	27.19	1171.35	3.32%	5.32%	0.07%
14	22789370	26.01	1171.01	23632083	27.70	1171.86	3.57%	6.12%	0.07%
16	23987166	26.34	1171.20	24925900	28.28	1172.09	3.77%	6.86%	0.08%
18	25137181	26.73	1171.22	26164504	28.91	1172.13	3.93%	7.55%	0.08%
20	26256711	27.17	1171.13	27366204	29.59	1172.07	4.05%	8.18%	0.08%

Table 12. Objective values comparison of two models when σ_z changes

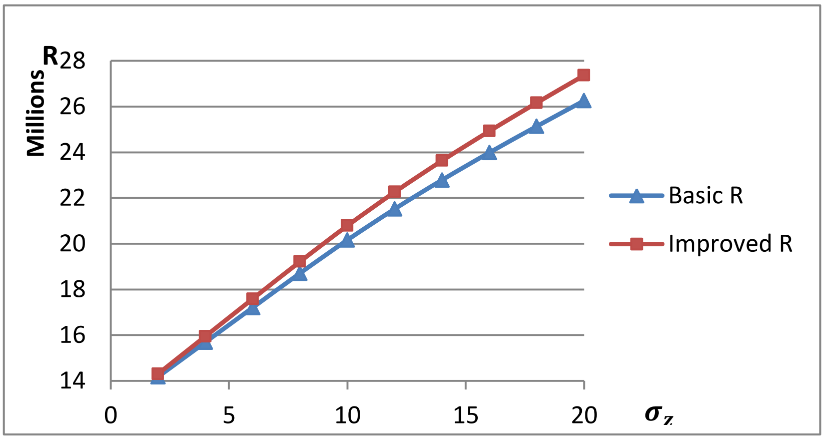

Figure 15. Total cost as σ_z increases

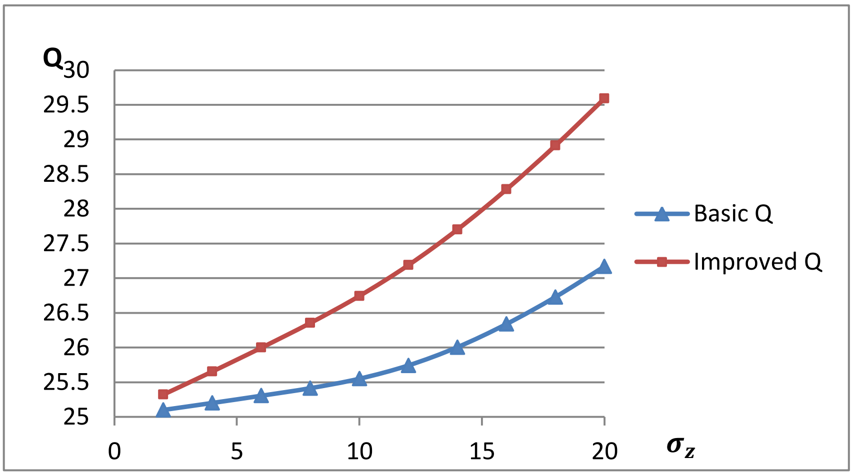

Figure 16. Optimal order quantity as σ_z increases

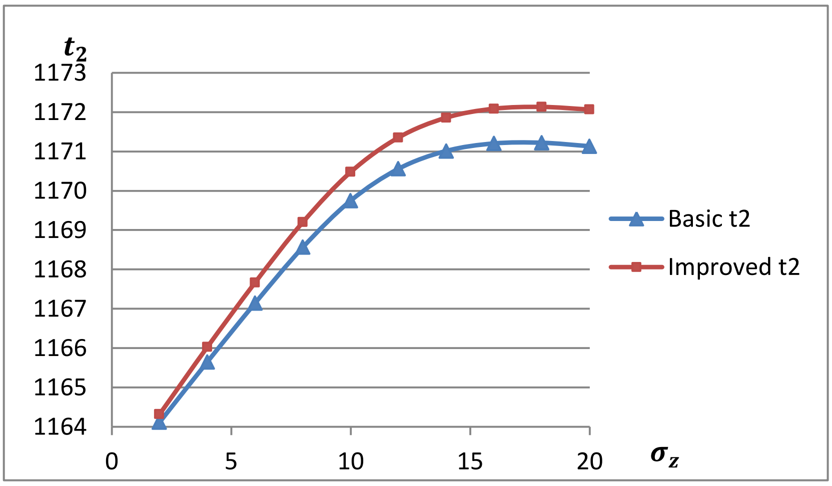

Figure 17. Optimal inventory replenishment time as σ_z increases

In the improved model, we consider when the exact Q^{th} failure happens, which is related to n, the total number of parts for a give PN in the observed fleet. When n increases, R, Q, and t_2 increase sharply at the beginning, followed by a flatter slope. The errors of R between the basic and improved models are in the range of 1.79% to 7.83% and the errors of Q between the two models are within 10%. As before, the errors of t_2 for the two models remain negligible, lower than 0.15%.

	Basic model			**Improved model**			**Percentage Error**		
n	R	Q	t_2	R	Q	t_2	R	Q	t_2
100	20161980	25.55	1169.75	20530166	26.38	1170.25	1.79%	3.14%	0.04%
200	20161980	25.55	1169.75	20787749	26.74	1170.48	3.01%	4.46%	0.06%
500	20161980	25.55	1169.75	21041153	27.12	1170.71	4.18%	5.78%	0.08%
1000	20161980	25.55	1169.75	21195563	27.35	1170.86	4.88%	6.58%	0.09%
5000	20161980	25.55	1169.75	21482789	27.79	1171.13	6.15%	8.05%	0.12%
10000	20161980	25.55	1169.75	21585992	27.95	1171.23	6.60%	8.57%	0.13%
100000	20161980	25.55	1169.75	21873606	28.39	1171.51	7.83%	9.99%	0.15%

Table 13. Objective values comparison of two models when n changes

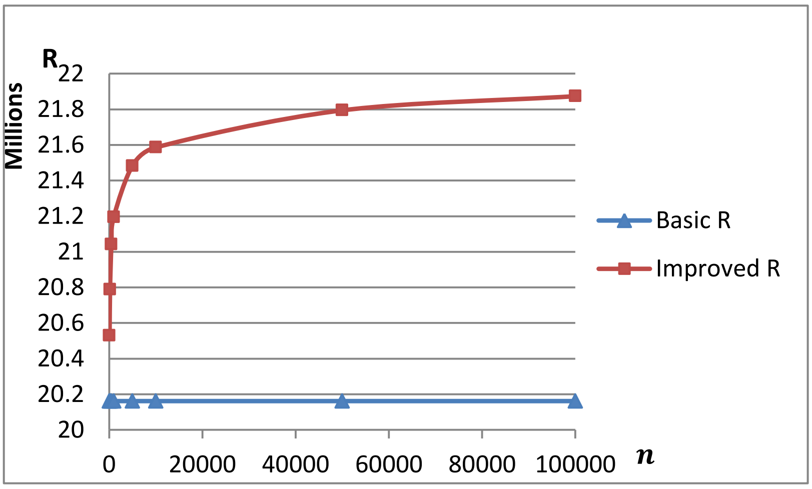

Figure 18. Total cost as n increases

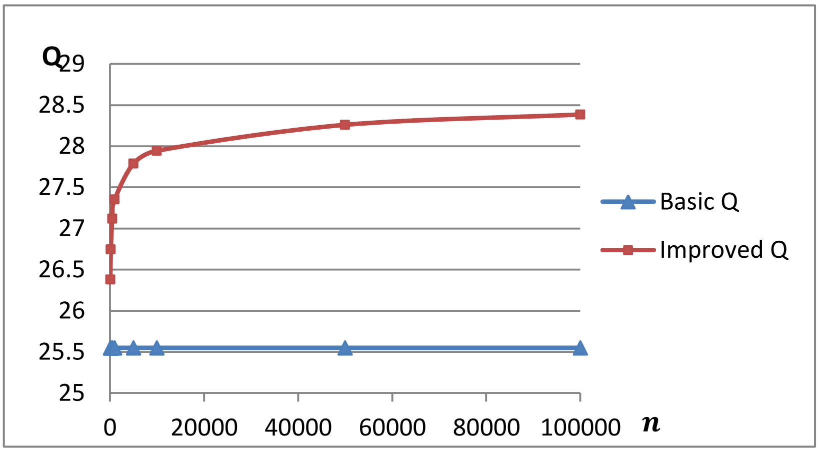

Figure 19. Optimal order quantity as n increases

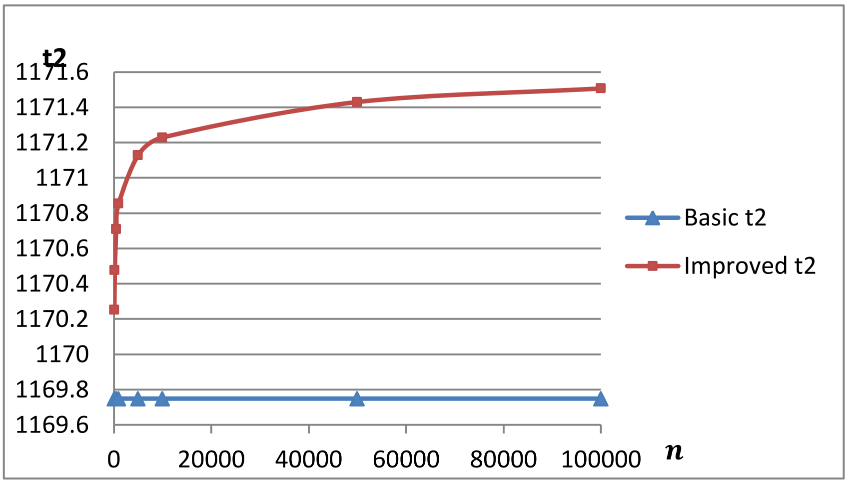

Figure 20. Optimal inventory replenishment time as n increases

2.7 Concluding Remarks

In this chapter, by using the airline industry as the background, we have developed two mathematical models to solve a single PN spare parts inventory management problem, where demands come from installed part failures. We aim to establish an efficient inventory policy that aims to minimize the total cost of stock outs and holding spare parts inventory.

Our models aim to reduce downtime due to spare parts shortage and excessive inventory holding cost by introducing parts failures distribution and triggering advanced orders at a proper time with a proper quantity. Compared with traditional forecast methods based on historical data, our models are more reasonable because they consider parts aging and focus on impending demands. Computational results indicate that our proposed basic inventory model can lead to a significant reduction in inventory cost, ranging from 5.68% to 98.03% in the first example, and 45.93% to 93.06% in the second example. We also observed that the values of t_2 in the two models always remain at the same level no matter how we change μ_x, σ_x, or μ_z, σ_z. Furthermore, compared with μ_x and μ_z, increased σ_x and σ_z will widen the gap between the two models, in terms of both the minimum total cost (R) and the optimal order quantity (Q). The most dramatic gap in Q appears when we change the value of σ_x, with a percentage error of 12.54%, while this also happens to R with a percentage error of 8.3% (See Table 10). In addition, as the improved model is concerned with the exact Q^{th} failure time, it introduces a new parameter n. When n is large enough, for example, 100 000, it may result in relatively large percentage errors of R and Q respectively between the two models. In our numerical analysis, both are controlled within 10% (See Table 13). The values of t_2 remain stable. An insight drawn from the numerical analyses is that the basic model can be used as a reasonable substitute for inventory planning due to its easier operation. Moreover, the calculation result of t_2 from the basic model can be used to approximate the result of the improved model which can effectively simplify the calculation process.

The models presented in this chapter are single period models with only one given part number and our numerical results are based on the normal distribution. It will be interesting to extend the model to a wider scope of applications, such as multiple periods, multiple given part numbers, budget limit, other parts failure distributions, to name a few.

CHAPTER 3: PROACTIVE SPARE PARTS PROCUREMENT INVENTORY MANAGEMENT FOR FAILURE-BASED MAINTENANCE POLICY -MULTIPLE PART NUMBER (PN) CASE

3.1 Introduction

In this chapter, we study multiple spare parts inventory policies. It is very common that there are multiple spare parts numbers in an inventory system and they face a budget constraint. We aim to establish an efficient spare parts inventory model for multiple PNs to achieve system optimization by considering budget issues and parts delivery time.

In practice, it is a common dilemma to maintain a balanced spare parts inventory level. It is understandable that plenty of spare parts in stock increases the service level but the inventory holding cost is usually very high; while a low stock level will reduce inventory holding cost, but may result in a low service level and lead to extremely high shortage cost. The more the types of spare parts are, the harder the problem can be solved. In our research, we assume that all demands are triggered by installed parts failures. Under an ideal situation, those parts should be in stock and in turn replenished by further procurement activities. However, an optimal inventory policy is often difficult to obtain due to demand uncertainty. Moreover, parts delivery time, budget limit, and a wide variety of PNs are all issues that may prevent decision-makers to derive reasonable inventory policies. Our models intend to incorporate all the aforesaid issues. We first introduce parts failure distributions in reliability engineering to predict impending demands. Then we use a multi-product newsvendor model to describe the total expected cost consisting of purchasing, holding, and shortage costs. Unlike the existing newsvendor models that focus on the determination of lot sizes, our proposed models consider the order timing as well. It is apparent that the total cost fluctuates with the lifetime and quantity of installed parts as the uncertainty in their failure influences the duration and numbers of spare parts shortage or overstock. Our proposed nonlinear programming (NLP) models can help spare parts managers to find optimal failure-based procurement inventory policies to minimize cost, with a limited budget constraint.

This chapter is organized as follows. In Section 3.2, we give a brief literature review in several related research streams. Then two mathematical models are put forward in Section 3.3: a basic mathematical model considering shortage period starts from mean time to failure (MTTF) and an improved mathematical model which takes into account the exact shortage time. Both models deal with multiple PNs with budget constraint. Solution methodologies are shown in Section 3.4. Three approaches are proposed with explanations of their advantages and disadvantages. Some numerical examples and calculation results are illustrated in Section 3.5. Experimental results demonstrate that the combination of a Lagrangian relaxation hheuristic and an iteration method is a reliable and stable approach to solving our NLP models. Section 3.6 furnishes a sensitivity analysis for three typical cases to examine the impact of important parameters on the optimal solution. Section 3.6 concludes the chapter and points out future research opportunities.

3.2 Literature Review

In some industries, such as the airline industry, the reorder point system (ROP) is widely used due to its easy operation. Future demand is often projected based on historical data. However, from a survey conducted by Ghobbar and Friend (2004), 152 out of 175 respondents were using the ROP system and about half were dissatisfied and considering implementing different systems. Therefore, finding more accurate and efficient forecasting methods seems to be a viable solution. Ghobbar and Friend (2003) discussed the forecasting of intermittent demand in relation to primary maintenance processes, and compared the experimental results of thirteen forecasting methods. Regattieri et al. (2005) analyzed the behavior of forecasting techniques when dealing with lumpy demand, and extended to compare twenty forecasting techniques. Both of the two papers found that the best approaches for intermittent demand are weighted moving average, Holt and Croston method. Though these approaches can improve forecast accuracy compared with ROP system, they ignore that in practice most spare parts demands are triggered by failures, which are highly related with operating hours, not past demands. This difference may lead great forecasting error. Our research realizes the correlation between part aging and failures, thus impending high demands might be forecasted even the past part demand is low, which is counter to the traditional forecasting system that past high demand triggers replenishment and past low demand scales back replenishment.

Parts failure distributions are widely used in joint optimization of spare parts inventory and maintenance. Some papers addressed the problems of spare parts provision based on failure-based repair policies. Simpson (1978) proposed an optimum solution structure for an n-period repairable inventory problem using a backward dynamic programming technique. Albright and Gupta (1993) modeled a two-echelon multi-indentured repairable-item inventory system where all failed modules would go to a single `depot' repair facility. Dhakar et al. (1994) presented a methodology to determine the optimal stocking levels for high-cost, low-demand, critical repairable spares based on an (S - 1, S) ordering policy. This stream of research focused solely on order quantity for repairable items. On the other hand, significant research is devoted to the joint optimization of spare parts inventory and age or block-based replacement policies. According to Vaughan (2005), age replacement refers to a scheme that a randomly failing item is replaced upon reaching some specified age T, or upon failure, whichever comes first. If we extend this policy to a group of items, it becomes block replacement. In certain circumstances, it is more economical to replace a group of items simultaneously rather than sequentially. This line of research started from one spare unit. Kaio and Osaki (1978) described an ordering policy with lead time for an operating unit in preventive maintenance (PM). This policy tries to find optimal regular ordering time to maximize the cost effectiveness. Thomas and Osaki (1978) extended the one unit model to find a proper unit operation and replacement time to minimize the total cost. Similar age replacement policy for one unit can also be found in Armstrong and Atkins (1996). The inventory policies of several identical spare units were developed based on previous research on one spare unit (see Kabir and Al-Olayan, 1994, 1996; Brezavscek and Hudoklin, 2003; Vaughan, 2005). These papers combined age or block replacement with periodic review spare-provisioning policies. Most recently, Wang (2012) presented a joint optimization approach for both the inventory control of the spare parts and the PM inspection interval. Enumeration and stochastic dynamic programming algorithms were used to optimize the order quantity, order intervals and PM intervals. The ordering policy therein is a typical (S, Q) policy in that ordering Q according to inventory level S at the time of ordering under a fixed order interval. To our knowledge, limited research deals with failure-based procurement inventory policies. Also, they typically handle single PN problems. In practice, however, multiple PNs in inventory must be managed from a systematic point of view. In addition decision-makers often have to consider some realistic constraints such as a limited budget.

The newsvendor problem has a rich history in operations management to determine optimal inventory levels (see Hadley and Within, 1963; Lau and Lau, 1995, 1996; Abdel-Malek et al., 2004, 2005, 2008; Zhang, 2010). These papers have analyzed multi-product newsvendor problems with constraints, thus have natural advantages in managing inventory with constraints from a system perspective. To the best of the author's knowledge, there is little research taking into account the application of multi-product newsvendor model with a budget constraint in spare parts procurement. Our research aims to bridge the gap by combining multi-product newsvendor model with installed parts failure distributions and a purchasing budget. Moreover, we help decision maker find when is the best time to order spare parts to realize cost minimization in each planning horizon.

In summary, this chapter combine maintenance data and inventory management to deal with a failure-based proactive procurement problem. We use installed parts failure distributions instead of historical demands to predict impending demands. Our inventory model is based on the multi-product newsvendor model which realizes system optimization with a budget constraint. In a single planning period, multiple types of spare parts are considered, and each type of spare parts is identified by a unique PN. Our inventory policy helps a planner to decide how many units to order and when to order to achieve total cost minimization, with a budget constraint. In our research, we assume only one purchase order is placed for each PN in the whole planning horizon. Different PNs can have different order quantities and order times.

3.3 Mathematical Model

3.3.1 Basic Mathematical Model

This research considers a generalized ordering policy for multiple PNs in a single period. For each type of parts i with a given PN, the length of the planning horizon is denoted by $\mathrm{T_i}$ $(0 \leq \mathrm{T_i} \leq \infty)$ and the order quantity in this period is denoted by $\mathrm{Q_i}$. The spare parts for replacement are delivered after a constant lead time $\mathrm{L_i}$. The demand is uncertain and depends on the parts failure distribution. We assume that the number of failures for parts i in period $\mathrm{T_i}$ follow a probability density function $g_i(\cdot)$ and cumulative distribution function $G_i(\cdot)$. The lifetimes of the operating parts are assumed independent with a probability density function $f_i(\cdot)$ and cumulative distribution function $F_i(\cdot)$. We also assume $g_i(\cdot)$ and $f_i(\cdot)$ are uncorrelated.

The following notations are used in the model formulation:

Indices

i= 1, …, m index of parts categories, where m is the total number of parts categories

Parameters

h_i the unit holding cost per unit time of product i

s_i the unit shortage cost per unit time of product i

T_i planning horizon of product i, can be infinite

L_i order lead time of product i

x_i lifetime for each part, a random variable

z_i demand quantity for each PN, a random variable

c_i unit price

K the budget limitation

$f_i(x_i)$ the PDF of failure distribution considering lifetime for each part

$g_i(z_i)$ the PDF of failure distribution considering number of failures for each PN

$F_i(x_i)$ the CDF of failure distribution considering lifetime for each part

$\mathrm{G_i}(z_i)$ the CDF of failure distribution considering number of failures for each PN

Define the following decision variables

$\mathrm{t}_{1\mathrm{i}}$ time point to issue order

$\mathrm{t}_{2\mathrm{i}}$ the parts arrival time

Q_{i} order quantity

The objective is to minimize the expected cost:

$$\text{Min R} = \sum_{i=1}^{m}\Big\{h_i(T_i - \mathrm{t}_{2\mathrm{i}})\int_0^{Q_i}(Q_i - \mathrm{z}_\mathrm{i})g_i(\mathrm{z}_\mathrm{i})d\mathrm{z}_\mathrm{i} + s_i\Big[T_i - \int_0^{\infty}x_i f_i(x_i)dx_i\Big]\int_{Q_i}^{\infty}(\mathrm{z}_\mathrm{i} - Q_i)g_i(\mathrm{z}_\mathrm{i})d\mathrm{z}_\mathrm{i} + \Big[h_i\int_{t_{2i}}^{\infty}(x_i - t_{2i})f_i(x_i)dx_i + s_i\int_0^{t_{2i}}(t_{2i} - x_i)f_i(x_i)dx_i\Big]Q_i + c_iQ_i\Big\}$$

(11)

Budget constraint:

$$\sum_{\mathrm{i}=1}^{\mathrm{m}} c_i Q_i \le \mathrm{K}$$

(12)

3.3.2 Improved Mathematical Model

The basic model presented above simplifies the function by assuming MTTF to be the Qth failure time to determine when the parts shortage period starts. In the improved mathematical model, we consider the exact Qth failure time to help us determine more accurate order quantity and order time. If we assume t_{Qi} is the Q_ith failure time for a given PN i. We also assume that the failure time of PN i is normally distributed, and follows $x_i \sim N(\mu_{xi}, {\sigma_{xi}}^2)$. n_i is the total number of parts in the observed fleet. Then equation (1) can be revised as:

$$\text{Min R} = \sum_{i=1}^{m}\Big\{h_i(T_i - \mathrm{t_{2i}})\int_0^{Q_i}(Q_i - \mathrm{z_i})g_{\boldsymbol{i}}(\mathrm{z_i})d\mathrm{z_i} + s_i\big(T_i - \mathrm{t}_{Qi}\big)\int_{Q_i}^{\infty}(\mathrm{z_i} - Q_i)g_{\boldsymbol{i}}(\mathrm{z_i})d\mathrm{z_i} + \Big[h_i\int_{t_{2i}}^{\infty}(x_i - t_{2i})f_i(x_i)dx_i + s_i\int_0^{t_{2i}}(t_{2i} - x_i)f_i(x_i)dx_i\Big]Q_i + c_iQ_i\Big\}$$

(13)

where the Q^{th} failure time is calculated by

$$\mathrm{t}_{Qi} = \sigma_{\mathrm{xi}} * F_S^{-1}\left(\frac{Q_i}{n_i}\right) + \mu_{\mathrm{xi}} \quad ,$$

(14)

with the same budget constraint:

$$\sum_{\mathrm{i=1}}^{\mathrm{m}} c_iQ_i \leq \mathrm{K}$$

(15)

3.4 Solution Methodology

3.4.1 GAMS and Its Solvers

GAMS and its solvers are firstly employed to solve our NLP models. Numerical experiments demonstrate that only CONOPT, MINOS, and MINOS 5 can generate solution results. Furthermore, MINOS 5 appears to be the most effective solver among the three in solving multiple PNs models. However, when the number of PNs is large, even MINOS 5 cannot always yield reliable results. Therefore, we have to find new ways to solve these multiple PNs models, especially for large-scale instances.

3.4.2 A Lagrangian Relaxation Heuristic

Zhang (2010) presented a Lagrangian relaxation heuristic to solve a mixed integer non-linear programming model for the multi-product newsvendor problem with both supplier quantity discounts and a budget constraint. The advantage of this approach is that it can relax the budget constraint and decompose the original problem into a series of sub-problems. It can also provide an error bound between the optimal solution and the approximate solution. Thus, for the improved model, we relax the budget constraint (15) and decompose function (13), then use GAMS and its solvers to solve these decomposed single PN problems.

By introducing a Lagrange multiplier λ, we relax the budget constraint by constructing the following Lagrangian dual problem.

$$\max_{\lambda} min\, L(Q_i, t_{2i}, \lambda) = R(Q_i, t_{2i}) - \lambda(K - \sum_{i=1}^{m} c_i Q_i)$$

(16)

The solution of the above Lagrangian dual problem, $\max_{\lambda} L(\lambda)$, gives the highest lower bound of the solution to the original problem. With a given value of Lagrangian multiplier λ, the lower bound $L(Q_i, t_{2i}, \lambda)$ can be rewritten as follows:

$$L\left(Q_i, t_{2i}, \lambda\right) = \sum_{i=1}^{m} R_i\left(Q_i, t_{2i}\right) - \lambda K$$

(17)

If we decompose the improved mathematical model into sub-problems, each for product i, we have:

$$Min\ R_i = h_i(T_i - t_{2i})\int_0^{Q_i}(Q_i - z_i)g_i(z_i)dz_i + s_i(T_i - t_{Qi})\int_{Q_i}^{\infty}(z_i - Q_i)g_i(z_i)dz_i + \left[h_i\int_{t_{2i}}^{\infty}(x_i - t_{2i})f_i(x_i)dx_i + s_i\int_0^{t_{2i}}(t_{2i} - x_i)f_i(x_i)dx_i\right]Q_i + c_iQ_i + \lambda c_iQ_i \tag{18}$$

where the Q^{th} failure time is calculated by

$$t_{Qi} = \sigma_{xi} * F_S^{-1}\left(\frac{Q_i}{n_i}\right) + \mu_{xi} \tag{19}$$

In order to obtain the highest lower bound of the Lagrangian dual problem, a sequence of Lagrangian multipliers are generated to repeatedly calculate the Lagrangian dual. At each iteration, the Lagrangian multiplier is updated by using a bisection method. The complete algorithm is described as follows:

Step 1. Set λ_L= 0, λ_U=10000 (to be adjusted).

Step 2. Let λ = (λ_L+ λ_U)/2.

Solve the sub problems with MINOS5.

LR = 0; /* Lagrange relaxation value*/

Loop (i,

Solve sub-problem using NLP minimizing R;

LR = LR + R).

Step 3: Update Lagrange relaxation value.

LR = LR − λ*K.

If new LR is larger than before, then it is improved.

(need to use a scalar to record previous LR)

Step 4: Update λ.

Let ca_gap = $\sum_{\mathrm{i}=1}^{\mathrm{m}} c_i Q_i - \mathrm{K}$,

If |ca_gap| < 10E-04 or | λ_L- λ_U |< 10E-05,

Stop.

If ca_gap < 0, then λ_U= λ;

Else λ_L= λ.

Step 5: Go to step 2.

3.4.3 The Combination of the Lagrangian Relaxation Heuristic and Iteration Method

The aforesaid two approaches use GAMS and its solvers to solve our NLP models. When the number of PNs is small, such as m=1 or m=2, both approaches can give reasonable results. However, when m is large, m≥10 for instance, solvers such as CONOPT, MINOS, MINOS 5 become unreliable. Excessive long time is usually needed to run the programme, and often, the final results are either unbounded or unreasonable. For example, the order quantity for some parts may be close to 0, or no feasible result. The part delivery time may approach infinity or no feasible result is obtained. Accordingly, it is essential to explore a new method that does not use GAMS solvers.

In Chapter 2, we noticed that the optimal parts arrival time t_2 for the basic and improved models are very close. Thus the iteration method is considered as an effective approach to solve our single PN NLP models. In this section, we continue assuming that all the observed parts failures are normally distributed, that is, for a given PN i, its part lifetime $f_i(x_i)$ follows a normal distribution with a mean of μ_{xi} and a standard deviation of σ_{xi}. Furthermore, the number of failures for the specified PN ($g_i(z_i)$) is normally distributed with a mean of μ_z and a standard deviation of σ_z. If we assume $t_{Qi} = \int_0^{\infty} x_i f_i(x_i) dx_i = \mu_{xi}$, the improved model will be simplified as the basic model. Lagrangian relaxation Function (8) can be rewritten as

$$Min\ R_i = h_i(T_i - t_{2i}) \int_0^{Q_i} (Q_i - z_i) g_i(z_i) dz_i + s_i(T_i - \mu_{xi}) \int_{Q_i}^{\infty} (z_i - Q_i) g_i(z_i) dz_i$$

$$+ \left[h_i \int_{t_{2i}}^{\infty} (x_i - t_{2i}) f_i(x_i) dx_i + s_i \int_0^{t_{2i}} (t_{2i} - x_i) f_i(x_i) dx_i \right] Q_i + c_i Q_i + \lambda c_i Q_i \tag{20}$$

There are only two decision variables in function (20), Q_i and t_{2i}. The optimal solution can be found by the following procedure:

$$\frac{\partial R_i}{\partial Q_i} = h_i(T_i - t_{2i}) G_i(Q_i) - s_i(T_i - t_{Qi})[1 - G_i(Q_i)] + \left[h_i \int_{t_{2i}}^{\infty} (x_i - t_{2i}) f_i(x_i) dx_i + s_i \int_0^{t_{2i}} (t_{2i} - x_i) f_i(x_i) dx_i \right] + c_i(1 + \lambda) \tag{21}$$

It follows that

$$\frac{\partial^2 R_i}{\partial {Q_i}^2} = \left[h_i(T_i - t_{2i}) + s_i(T_i - t_{Qi}) \right] g_i(Q_i) \geq 0, \text{for all } Q_i \geq 0.$$

Because the second derivative is nonnegative, the function $R_i(Q_i)$ is said to be convex. The optimal solution, denoted by $Q_i{*}$, occurs where $\frac{\partial R_i}{\partial Q_i}$ equals zero. That is,

$$G_{\boldsymbol{i}}(Q_i{}^*) = \frac{s_i(T_i - \mathrm{t}_{Qi}) - \left[h_i \int_{t_{2i}}^{\infty}(x_i - t_{2i}) f_i(x_i) dx_i + s_i \int_0^{t_{2i}}(t_{2i} - x_i) f_i(x_i) dx_i\right] - c_i(1+\lambda)}{h_i(T_i - t_{2i}) + s_i(T_i - \mathrm{t}_{Qi})}$$

(22)

Also,

$$\frac{\partial R_i}{\partial t_{2i}} = -h_i \int_0^{Q_i}(Q_i - z_i) g_{\boldsymbol{i}}(z_i) dz_i - Q_i h_i[1 - F_i(t_{2i})] + Q_i s_i F_i(t_{2i})$$

(23)

It follows that

$$\frac{\partial^2 R_i}{\partial t_{2i}{}^2} = Q_i f_i(t_{2i})(h_i + s_i) \geq 0, \text{for all } Q_i \geq 0.$$

Because the second derivative is nonnegative, the function R_i (t_{2i}) is said to be convex. The optimal solution, say $t_{2i}{*}$, occurs where $\frac{\partial R_i}{\partial t_{2i}}$ equals zero. That is,

$$F_i(t_{2i}{}^*) = \frac{h_i \int_0^{Q_i}(Q_i - z_i) g_{\boldsymbol{i}}(z_i) dz_i + Q_i h_i}{Q_i(h_i + s_i)}$$

(24)

From equation (12), we have

$$Q_i{}^* = G_{\boldsymbol{i}}{}^{-1}\left\{\frac{s_i(T_i - \mathrm{t}_{Qi}) - h_i(\mu_{\mathrm{xi}} - t_{2i}) - (h_i + s_i)\left[(t_{2i} - \mu_{\mathrm{xi}}) F_s\left(\frac{t_{2i} - \mu_{\mathrm{xi}}}{\sigma_{\mathrm{xi}}}\right) + \sigma_{\mathrm{xi}} f_s\left(\frac{t_{2i} - \mu_{\mathrm{xi}}}{\sigma_{\mathrm{xi}}}\right)\right] - c_i((1+\lambda))}{h_i(T_i - t_{2i}) + s_i(T_i - \mathrm{t}_{Qi})}\right\}$$

(25)

Here $f_s(\cdot)$ is the standard normal density function, and $F_s(\cdot)$ is the standardized cumulative distribution function for normal density function $f_s(\cdot)$. $G_{\boldsymbol{i}}{}^{-1}(\cdot)$ is the inverse of the cumulative density function for normal density function $g_i(\cdot)$.

From equation (14), we have

$$t_{2i}{}^* = F_i{}^{-1}\left\{\frac{h_i(Q_i - \mu_{\mathrm{zi}}) G_s\left[\frac{(Q_i - \mu_{\mathrm{zi}})}{\sigma_{\mathrm{zi}}}\right] + h_i \sigma_{\mathrm{zi}} g_s\left[\frac{(Q_i - \mu_{\mathrm{zi}})}{\sigma_{\mathrm{zi}}}\right] + Q_i h_i}{Q_i(h_i + s_i)}\right\}$$

(26)

Here $g_s(\cdot)$ is the standard normal density function, and $G_s(\cdot)$ is the standardized cumulative distribution function for normal density function $g(z)$. ${F_i}^{-1}(\cdot)$ is the inverse of the cumulative density function for normal density function $f_i(\cdot)$.

Because the closed forms of ${Q_i}^*$ and ${t_{2i}}^*$ are hard to be determined from function (25) and (26), we develop a numerical algorithm to find solutions. This procedure converges in a finite number of iterations. The steps of the algorithm are

Step 1. Set λ_L = 0, λ_U= 10000 (to be adjusted).

Step 2. Let λ = (λ_L+ λ_U)/2.

Step 3. Set $t_{Q_i} = \mu_{\mathrm{xi}}$, $t_{2_i} = 0$.

Step 4. Use the function (25) to find Q_i.

Step 5. Use Q_i to update t_{2_i} by function (26).

Step 6. Update t_{Q_i} by function (19).

Step 7. If $|t_{Q_i} - t_{Q_i last}|$< 1.0E-8 the optimal solution is ${Q_i}^* = Q_i$, ${t_{2_i}}^* = t_{2_i}$, and

${t_{Q_i}}^* = t_{Q_i}$. Otherwise, go to step 4.

Step 8. Calculate R_i by the RHS of function (18).

Step 9. Calculate the Lagrangian relaxation value LR =$\sum_{i=1}^{m} R_i$.

Step 10: Update the Lagrangian relaxation value.

LR = LR – λ*K.

If new LR is larger, then it is improved.

(need to use a scalar to record previous LR)

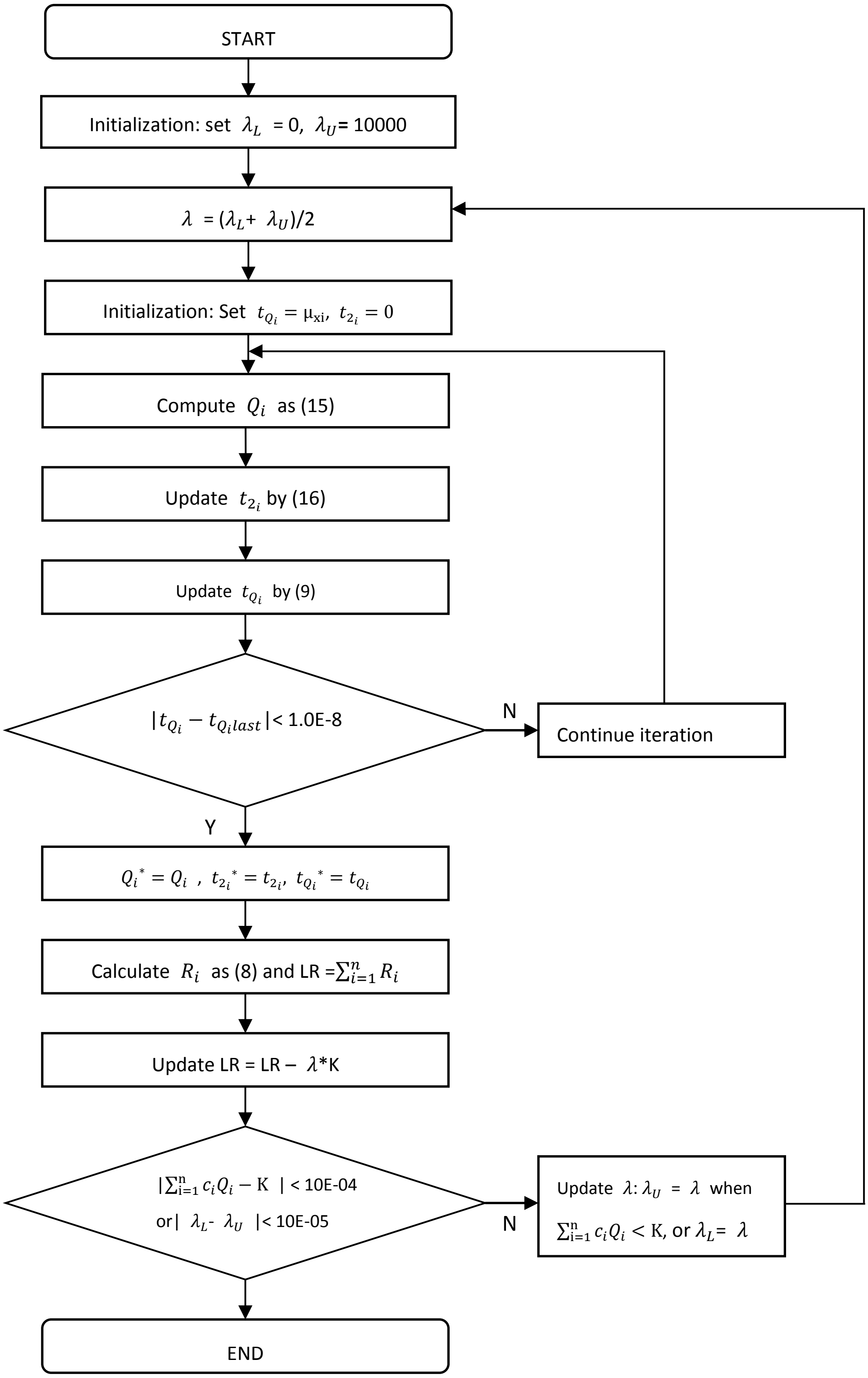

Figure 21. The flowchart of the combination of Lagrangian heuristic and iteration method

Step 11: Update λ.

Let ca_gap = $\sum_{i=1}^{m} c_i Q_i - K$,

If |ca_gap| < 1.0E-04 or | λ_L-λ_U|< 1.0E-05,

Stop.

If ca_gap < 0, then $\lambda_U = \lambda$;

Else $\lambda_L = \lambda$.

Step 12: Go to step 2.

The flowchart of the algorithm is illustrated in Figure 21.

Once we know $Q_i^*$and t_{2i}^* , the optimal time to issue order for a given PN index i can be found by $t_{1i} = t_{2i}^* - L_i$.

3.5 Numerical Examples and Results

In this section, we give some numerical examples and compare their results for the three methods. As mentioned before, we still assume parts demands follow normal distributions.

We first consider an example with m=2 PNs to test our Lagrangian relaxation approach and the Iteration method. Parameter values are listed in Table 14 that are adapted from single PN examples in Chapter 2.

	c(i)	h (i)	s(i)	μx(i) days	σx(i) days	T(i) days	μz(i)	σz(i)	n(i)	K
part i=1	449586	307.94	6158.71	243.6	65.9	1825	25	10	200	2.32E+07
part i=2	449586	615.87	2463.48	1218	65.9	1825	25	10	200	2.32E+07

Table 14. Parameter values for two PNs example

The calculation results are summarized in Table 15 for three different methods: GAMS and its solver MINOS 5, the combination of the Lagrangian relaxation heuristic and GAMS /MINOS 5, as well as the combination of Lagrangian relaxation heuristic and iteration method**.** Compared with single PN examples in Chapter 2, we can see that GAMS/MINOS 5 produces good and accurate results in this small size problem, but it does not work well once it is combined with the Lagrangian relaxation method. However, if the iteration method instead of the GAMS solver is applied to these relaxed and decomposed sub-problems, the optimal solutions are almost the same as GAMS/MINOS 5 results. The tiny difference may be caused by numerical errors. Accordingly, two conclusions can be drawn based on this example: the GAMS/MINOS 5 solver is not reliable once we combine it with the Lagrangian relaxation heuristic due to the complexity of our NLP models. Furthermore, the combination of Lagrangian relaxation heuristic and iteration method is feasible and effective in solving these proposed NLP models.

		Q	t_2	t_Q	R
GAMS/MINOS5	part i=1	34.53	142.12	181.41	5.64E+07
	part i=2	17.00	1,165.46	1,127.56	
Lagrangian relaxation &GAMS solver	part i=1	0	5121955	1.00E+12	-2.75E+16
	part i=2	0	5121955	2.00E+12	
Lagrangian relaxation & Iteration	part i=1	34.64	140.58	181.55	5.64E+07
	part i=2	16.89	1,158.52	1,127.33	

Table 15. Calculation results and comparison for two PNs example

	GAMS/MINOS5				Lagrangian relaxation & Iteration			
	Q	t_2	t_Q	R	Q	t_2	t_Q	R
part i=1	0.02	/	21.60	1.47E+10	139.52	84.27	121.81	8.08E+09
part i=2	155.37	219.96	336.83		169.06	191.14	347.22	
part i=3	162.56	109.71	152.65		174.03	101.64	156.08	
part i=4	89.54	282.08	556.17		96.41	300.97	573.52	
part i=5	29.71	/	231.77		32.41	130.01	238.07	
part i=6	66.93	129.79	/		32.41	123.57	220.49	
part i=7	339.23	410.73	4.04E+11		182.96	389.90	614.96	
part i=8	152.25	427.50	692.77		163.74	425.97	707.14	
part i=9	20.27	502.41	673.26		22.46	478.61	689.63	
part i=10	183.00	224.27	372.25		198.42	223.82	382.03	

Table 16. Calculation results and comparison for the problem with m=10

Next, an example with m=10 PNs is investigated. After coding with GAMS, all the related parameters are generated randomly and automatically. The demands/ failures of each PN follow a normal distribution with $\sigma_z(i)$ from 3 to 50, and $\mu_z(i)$ from 3 to 6 times of $\sigma_z(i)$. The part lifetime for each PN is set to be a normal distribution with $\sigma_x(i)$ from 8 to 200, and $\mu_x(i)$ from 3 to 6 times of $\sigma_x(i)$. The unit price for each PN is generated randomly from 100 to 60000, with unit holding cost from 10% to 30% of the unit price, and unit shortage cost from 3 to 5 times of the unit value. The budget is calculated by a random number between 1.1 and 1.3 multiplying the expected total purchase cost. The computational results are reported in Table 16. From the incomplete solutions for PN 1, 5, and 6 on the left side of the table, we can see that GAMS/MINOS5 is not appropriate to process large numbers of PNs in our NLP models. On the other hand, the combination of the Lagrangian relaxation and iteration methods can give reasonable solutions for every PN. Moreover, the gap between the Lagrangian relaxation value and objective value is 0, and the minimum total cost on the right hand side is obviously lower than that on the left. Therefore, the combination of Lagrangian relaxation and iteration has its superiority in solving our proposed multiple PNs NLP models compared with GAMS/MINOS5.

To test the performance and stability of the integrating Lagrangian relaxation and iteration approach in solving large size problems, six instances are randomly generated with the number of PNs (m) ranging from 5 to 2000 with the other parameters such as c(i), h(i), s(i), $\mu_x(i)$, $\mu_z(i)$ remaining the same in GAMS codes as mentioned in the previous paragraph. The computation results in Table 17 indicate that the integrative Lagrangian relaxation and iteration approach reports extremely good solutions in terms of both solution quality and computing time: the gap between the Lagrangian relaxation value and objective solution value for all the instances are 0. For cases involving up to 200 PNs, the optimal solutions can be found in less than 1 second. Even for large scale cases with thousands of PNs, this approach can obtain results in less than half a minute.

m	Bound	Solution	GAP(%)	CPU Times (second)
5	2.90E+09	2.90E+09	0	0.031
10	8.08E+09	8.08E+09	0	0.031
20	1.20E+10	1.20E+10	0	0.062
200	1.24E+11	1.24E+11	0	0.719
1000	7.38E+11	7.38E+11	0	7.063
2000	1.37E+12	1.37E+12	0	21.391

Table 17. Test Problem sizes, solutions, relative gap, and running time

3.6 Sensitivity Analysis

In order to understand how important parameters affect decision variables, we conduct a sensitivity analysis with the improved multiple PNs model for three different cases, which are typical representations based on our earlier analysis on single PN models in Chapter 2. The basic input parameter values are generated randomly in GAMS and are listed in Table 18. We choose a small size example, m=3, because it is easy to operate and clear to investigate. In each case, we only change one parameter value at a time while the others are kept the same.

	c(i)	h (i)	s(i)	μx(i) days	σx(i) days	T(i) days	μz(i)	σz(i)	n(i)	K
part i=1	10387.65	1664.39	38430.80	621.48	104.04	1825	165.56	33.07	367.55	
part i=2	50611.68	8019.04	238509.58	1055.44	199.64	1825	45.20	10.50	109.33	7197117.52
part i=3	33067.48	4788.52	103641.02	404.05	119.12	1825	60.18	14.75	150.31	

Table 18. Basic values of parameters

K		5997598	6597358	7197118	7796877	8396637	8996397
	part i=1	167.715	181.792	195.932	209.967	220.938	220.938
Q	part i=2	44.055	48.959	53.834	58.786	62.989	62.989
	part i=3	61.261	67.47	73.704	79.854	84.588	84.588
	part i=1	441.573	444.137	447.081	449.874	451.785	451.785
t_2	part i=2	686.873	692.081	698.376	704.546	709.016	709.016
	part i=3	201.928	205.395	209.23	212.755	215.095	215.095
	part i=1	610.056	620.068	630.11	640.158	648.132	648.132
t_Q	part i=2	1006.393	1029.256	1051.645	1074.337	1093.782	1093.782
	part i=3	376.203	388.747	401.172	413.399	422.871	422.871
R	SUM	2.72E+09	1.75E+09	1.23E+09	1.02E+09	1.01E+09	1.01E+09

Table 19. Decision variables and objective value change when K increases

First, we observe how the budget limit (K) influences our decision variables and objective function value. The test data are summarized in Table 19. K is changed from 5 997 598 to 8 996 397, which is actually from the value of $\sum_{i=1}^{m} c_i\mu_{zi}$ to $1.5\sum_{i=1}^{m} c_i\mu_{zi}$. We can see that Q, t_2, and t_Q increase as budget limit K is increased. While the total expected cost R decreases when a higher budget is allowed due to the reduction of shortage cost. The lowest optimal value is obtained at about K=8 396 637 (which is $1.4\sum_{i=1}^{m} c_i\mu_{zi}$), and after that point the decision variables and objective value do not change any more even if the budget limit is further increased. It is due to the reason that the budget limit is likely a binding constraint when K < 8 396 637 and becomes redundant once K ≥ 8 396 637.

Figure 22 describes that the order quantities Q for all the three PNs gradually increase with similar growth trends as the value of K increases. Figure 23 shows that the part arrival times t_2 for all the three PNs slightly increase in K. The expected total cost curve in Figure 24 reveals that initial budget limit increase will significantly reduce the expected total cost due to shortage cost savings. However, the expected total cost will not decrease any more once the budget constraint becomes nonbinding.

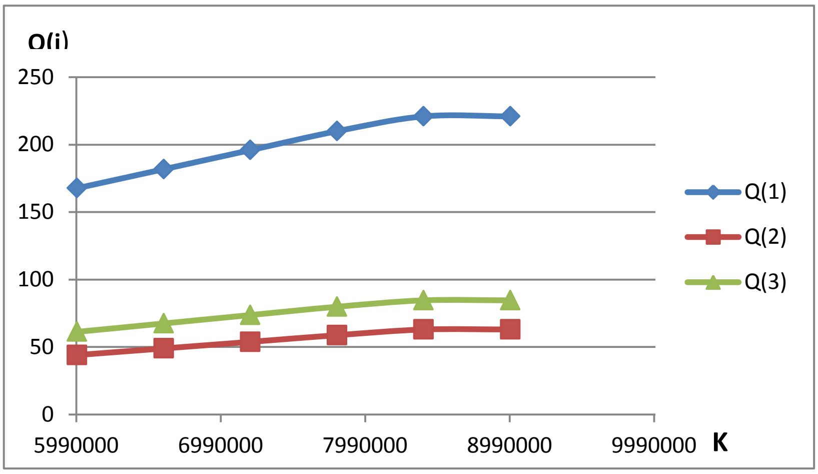

Figure 22. Order quantities change when K increases

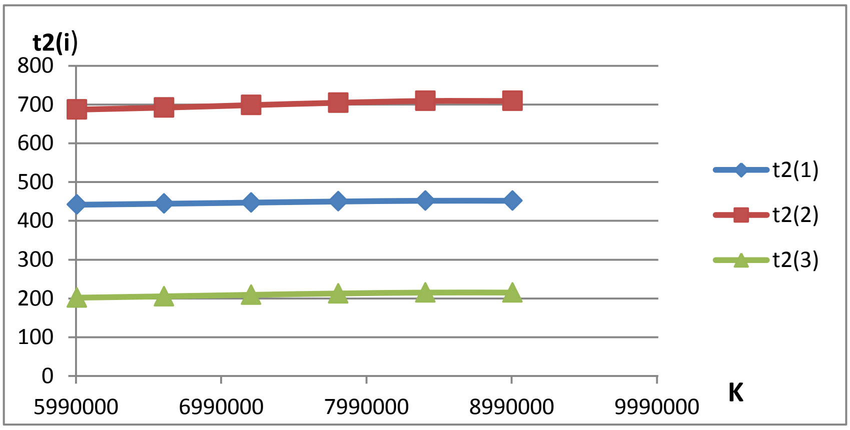

Figure 23. Parts arrival times change when K increases

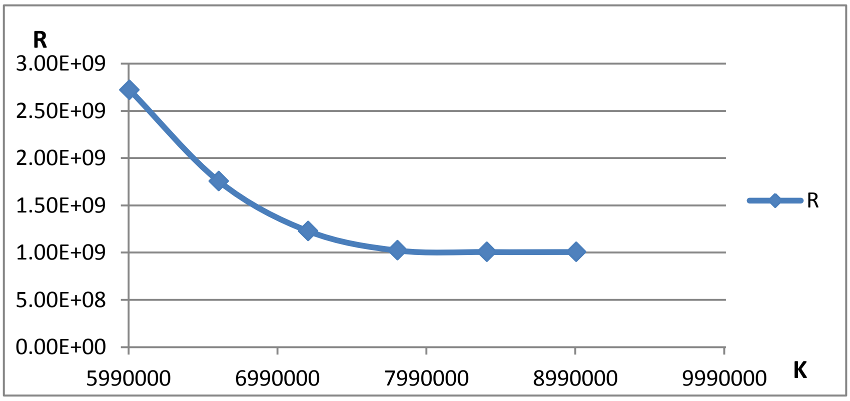

Figure 24. Expected total cost changes when K increases

Next, we investigate how the decision variables and objective function value change as the MTTF (μ_{x1}) for PN 1 increases. When K=7 197 118 (which is $1.2\sum_{i=1}^{m} c_i\mu_{zi}$), the optimal replenishment plans for all the three PNs are summarized in Table 20, and are illustrated in Figure 25, 26, and 27. When μ_{x1} increases, we can see that the part arrival times for PN 1 are postponed due to their longer expected lifetimes. The part shortage period for PN 1 also starts later due to the same reason, however, the values of t_2 and t_Q for the other two parts almost remain at the same levels. Order quantities for the three parts are also kept more or less the same as a result of budget constraint which restricts decision-makers' purchase of parts. Figure 27 shows that the expected total cost consistently decreases as μ_{x1} increases, which is likely due to the shortage cost savings for PN 1.

μ_{x1}		312.12	364.14	416.16	468.18	520.2	572.22	621.475
	part i=1	198.83	198.398	197.946	197.472	196.975	196.453	195.932
Q	part i=2	53.509	53.558	53.608	53.661	53.717	53.776	53.834
	part i=3	73.291	73.353	73.417	73.485	73.555	73.63	73.704
	part i=1	138.327	190.258	242.185	294.107	346.024	397.936	447.081
$\mathbf{t_2}$	part i=2	697.951	698.015	698.081	698.15	698.223	698.3	698.376
	part i=3	208.979	209.016	209.055	209.097	209.14	209.185	209.23
	part i=1	322.82	374.532	426.23	477.912	529.578	581.226	630.11
$\mathbf{t_Q}$	part i=2	1050.157	1050.378	1050.61	1050.853	1051.108	1051.376	1051.645
	part i=3	400.351	400.474	400.602	400.736	400.877	401.025	401.172
R	SUM	1.28E+09	1.27E+09	1.26E+09	1.25E+09	1.24E+09	1.24E+09	1.23E+09

Table 20. Decision variables and objective value change when μ_{x1} increases

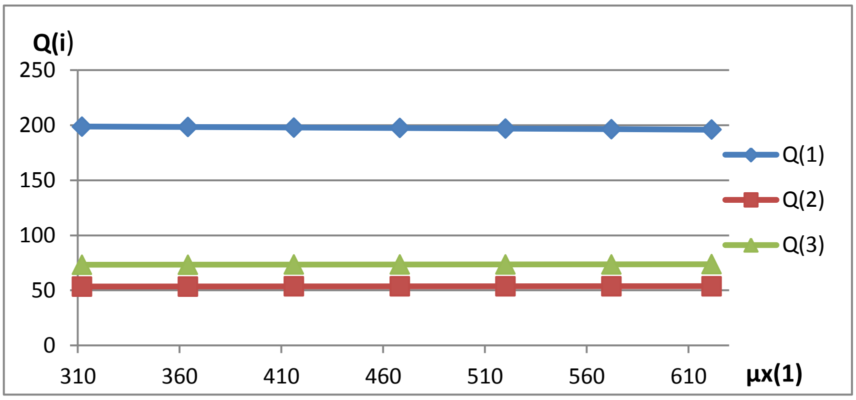

Figure 25. Order quantities change when μ_{x1} increases

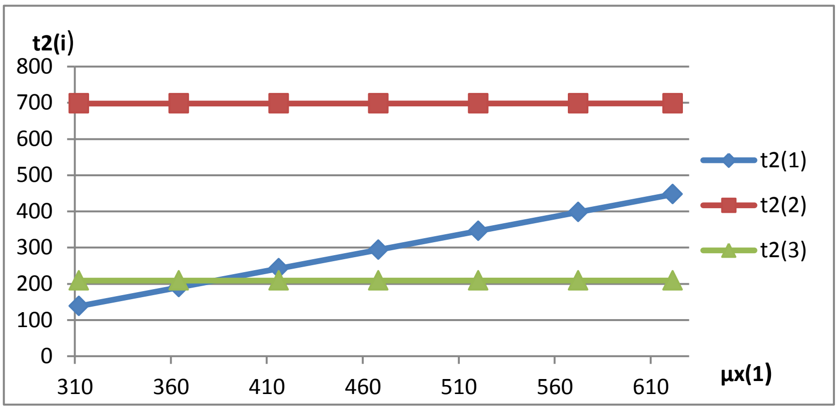

Figure 26. Parts arrival times change when μ_{x1} increases

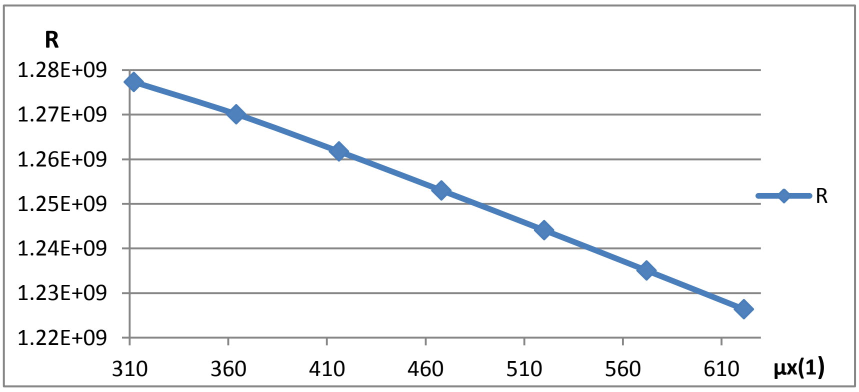

Figure 27. Expected total cost changes when μ_{x1} increases

Finally, we examine the impact of changing the standard deviation of the first PN (σ_{x1}). The numerical results are shown in Table 21. When σ_{x1} increases, lifetime uncertainty for the first PN increases. Thus, for this given PN, more parts are required in order to reduce the risk of shortage. However, as Figure 28 shows, the order quantity roughly remains at the same level, likely due to budget limit. The only thing that the decision-maker can adjust is replenishing PN 2 earlier to handle the challenges of part shortage, as reflected in Figure 29. Although this adjustment helps to alleviate shortage risk, it will nevertheless drive up inventory cost as depicted in Figure 30, where the expected total cost gradually grows as σ_{x1} increases. This increased cost is probably incurred by the higher inventory holding cost as well as higher risk of part shortage due to increased uncertainty of first PN lifetime.

σ_{x1}		104.04	130	160	190	215
	part i=1	195.932	196.327	196.183	196.04	195.921
Q	part i=2	53.834	53.79	53.806	53.822	53.835
	part i=3	73.704	73.648	73.668	73.689	73.706
	part i=1	447.081	354.416	304.107	253.781	211.83
t_2	part i=2	698.376	698.318	698.339	698.36	698.378
	part i=3	209.23	209.196	209.208	209.221	209.231
	part i=1	630.11	583.361	585.774	588.13	590.048
t_Q	part i=2	1051.645	1051.441	1051.515	1051.588	1051.649
	part i=3	401.172	401.061	401.101	401.142	401.176
R	SUM	1.23E+09	1.26E+09	1.28E+09	1.30E+09	1.31E+09

Table 21. Decision variables and objective value change when σ_{x1} increases

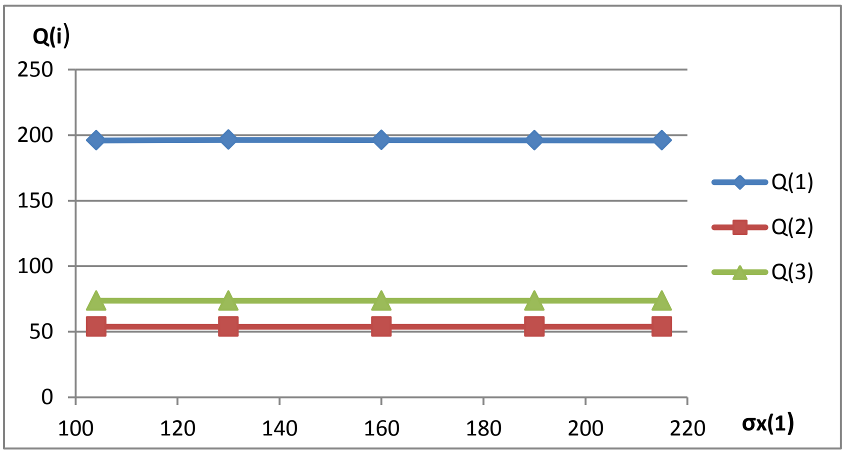

Figure 28. Order quantities change when σ_{x1} increases

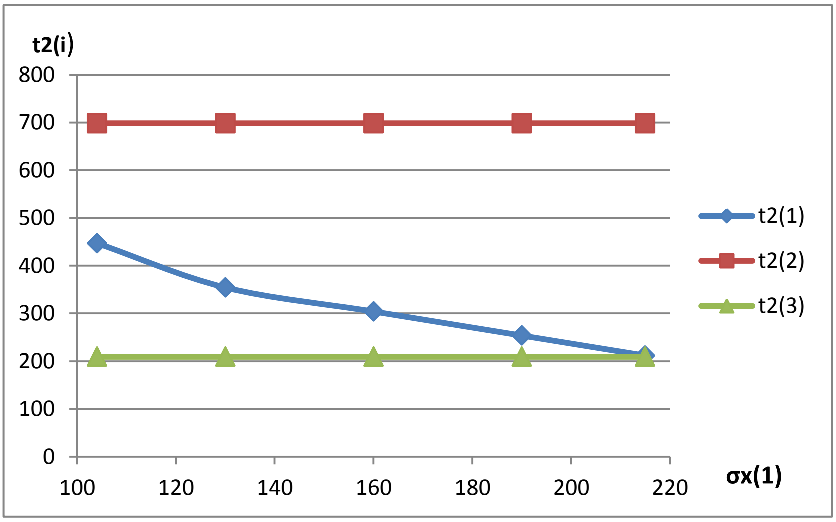

Figure 29. Parts arrival times change when σ_{x1} increases

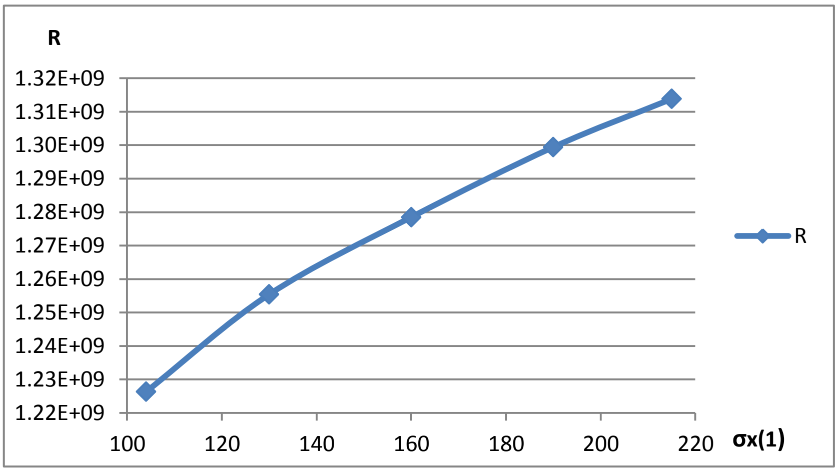

Figure 30. Expected total cost changes when σ_{x1} increases

3.7 Conclusion and Future Work

In this chapter, we have developed two mathematical models to manage failure-based spare parts procurement problems, where demands arise from installed parts failures. In the beginning of each planning horizon, spare parts managers have a limited budget on hand and they have to make procurement plans. Generally speaking, they have to face the following two problems: What is the order quantity for each PN? When to place a purchase order for each given PN? These two questions need to be addressed with the objective of minimizing the total cost of purchasing, holding and shortage within the available resources. Our research attempts to answer these questions by establishing an efficient inventory policy. The aim is to achieve a balance between the inventory holding and shortage cost. We also consider part delivery time in our models which help decision-makers place orders in advance, therefore reduce downtime and unnecessary cost. Compared with traditional forecasting methods based on historical data, our models address demand for spare parts by considering parts aging and wear-out. Part failures distributions are introduced to predict impending demands. The proposed multiple PN NLP models cannot be properly handled by GAMS and its solvers due to the complexity of the model, especially for large size cases. Therefore, we explore a Lagrangian relaxation heuristic method to relax the budget constraint and decompose the large size problem into multiple single PN sub-problems. To overcome the instability of GAMS solvers, we further combine an iteration method with the Lagrangian relaxation heuristic, and develop an integrative algorithm to find the optimal solution. Both small size and large size numerical examples show that the gap between the Lagrangian relaxation value and objective function value is 0. Our proposed NLP models and solution methodologies can greatly help decision-makers to find optimal order quantities and order time from a total cost minimization perspective under a limited budget constraint. In addition, our inventory models can also be applied to extremely large-scale problems. For example, even we have 2000 PNs in the inventory system, the solution still can be found expeditiously within 21.4 seconds in our numerical experiment.

The models presented in this chapter are single period models with multiple PNs under a budget constraint. Our numerical examples are based on normal distributions and the integrative approach of the Lagrangian relaxation and iteration method proves to be effective in solving the NLP models for both large and small numbers of PNs. It will be interesting to extend the problem to a wider scope of applications, such as multiple periods, other parts failure distributions, spare parts procurement based on both failures and PM, purchase plan considering repairable items, to name a few.

CHAPTER 4: CONCLUSIONS AND FUTURE RESEARCH

This thesis developed a series of non-linear programming models to obtain optimal spare parts replenishment policies for maintenance use due to failures replacements. In consideration of parts lead time, purchase orders must be issued in advance to reduce downtime, especially for those parts with long delivery time and cumbersome customs clearance processes. We take into account reliability issues and introduce part failure distributions from two perspectives: failure time and failure numbers. Therefore, part demands or part failures can be predicted based on two factors: time and quantity. Compared with traditional inventory policies, our proactive and efficient spare parts inventory policies not only consider the optimal order quantity, but also take into account the optimal order timing in a single period. Thus, advance orders can be triggered at appropriate time and the total cost, consisting of purchase, holding and shortage cost, can be minimized.

We first introduce two single PN models, the basic and improved models, in a single period. Our numerical results are based on normal distributions. Compared with traditional forecast methods based on historical data, our models seem more reasonable because they consider part aging and its impact on demands. Computational results indicate that our proposed basic inventory model can lead to a significant reduction in inventory cost, ranging from 5.68% to 98.03% in the first example, and 45.93% to 93.06% in the second example. We also found that the basic model can be used as a reasonable substitute to carry out inventory planning due to its ease for use. Moreover, the computation results of part arrival time from the basic model can be employed to approximate the result of the improved model. Although the spare parts replenishment policies in this research are mainly focused on aircraft maintenance, the models are put in a generic framework and can be used in other industries with similar concerns.

In Chapter 3, we extend these models to the multiple PNs case with a budget constraint. Due to the complexity of the two NLP models, we have explored several approaches to find a reliable and reasonable solution. Because GAMS and its solvers are not applicable to solving large-scale instances, especially out complex NLP models, we introduce a Lagrangian relaxation heuristic method to relax the budget constraint and decompose the large size problem into multiple single PN sub-problems. The advantage of the Lagrangian relaxation heuristic is that it can provide a measure for the gap between the optimal solution and the approximate solution. To overcome the instability of solvers in GAMS, we further integrate an iteration method into the Lagrangian relaxation heuristic, and develop an iterative algorithm to find the optimal solution. Numerical experiments demonstrate that the gap between the Lagrangian relaxation and the objective function value is 0 for both small size and large size examples. In summary, our proposed NLP models and solution methodologies are able to help decision makers find optimal order quantities and order time to minimize the total cost for multiple PNs under a limited budget constraint. As illustrated by our numerical examples, the proposed inventory models and the integrative solution procedure can be applied to large-scale problems. For the example with 2000 PNs, the solution still can be expeditiously obtained within 21.4 seconds.

The models presented in this thesis have addressed single PN and multiple PNs spare parts inventory management problems. Current models are limited to a single period and numerical examples are carried out based on normal distributions to characterize part failure time and the number of failures. It will be interesting to extend the models along a number of directions, such as multiple periods, other parts failure distributions, spare parts procurement based on both failures and PM, and purchase plan considering repairable items.

REFERENCES

Abdel-Malek, L., Montanari, R., Morales, L.C., 2004. Exact, approximate, and generic iterative models for the multi-product newsboy problem with budget constraint. International Journal of Production Economics, 91 (2), 189–198.

Abdel-Malek, L., Montanari, R., 2005. An analysis of the multi-product newsboy problem with a budget constraint. International Journal of Production Economics, 97, 296–307.

Abdel-Malek, L., Montanari, R., Meneghetti, D., 2008. The capacitated newsboy problem with random yield: the Gardener problem. International Journal of Production Economics, 115 (1), 113–127.

Albright, S.C., Gupta, A., 1993. Steady-state approximation of a multi-echelon multi-indentured repairable-item inventory system with a single repair facility. Naval Res. Log.,40, 479–493.

Allen, S.G., D'esopo, D.A., 1968. An ordering policy for repairable stock items. Operation Research, 16 (3), 482–489.

Armstrong, M., Atkins, D., 1996. Joint optimization of maintenance and inventory policies for a simple system. IIE Transactions, 28 (5), 415–424.

Bacchetti, A., Saccani, N., 2012. Spare parts classification and demand forecasting for stock control: investigating the gap between research and practice. Omega, 40, 722-737.

Batchoun, P., Ferland, J.A. and Cléroux, R., 2003. Allotment of aircraft spare parts using genetic algorithms. Pesquisa Operacional, 23(1), 141-159.

Brezavscek, A., Hudoklin, A., 2003. Joint optimization of block-replacement and periodic-review spare-provisioning policy. IEEE Transactions on Reliability, 52 (1), 112–117.

Byington, C., Roemer, M.J., Galie, T., 2002. Prognostic enhancements to diagnostic systems for improved condition-based maintenance. Proceedings of IEEE Aerospace Conference, vol. 6, Big Sky, USA, March 9–16, pp. 2815–2824.

Campbell, H.S., 1963. The Relationship of Resource Demands to Air Base Operations. RAND, Santa Monica.

Chopra, S., Meindl, P., 2010. *Supply Chain Management: Strategy, Planning, and Operation.* New Jersey: Pearson.

Deshpande, V., Iyer, A. V., Cho, R., 2006. Efficient supply chain management at the U.S Coast Guard using part-age dependent supply replenishment policies. Operation Research, 54 (6), 1028-1040

deSmidt-Destombes, K.S., van der Heijden, M.C., van Harten, A., 2007. Availability of k-out-of-N systems under block replacement sharing limited spares and repair capacity. International Journal of Production Economicse, 107, 404–421

de Smidt-Destombes, K.S., van der Heijden, M.C., van Harten, A., 2009. Joint optimisation of spare part inventory, maintenance frequency and repair capacity for k-out-of-N systems. International Journal of Production Economics, 118, 260–268.

Dhakar, T., Schmidt, C., Miller, D., 1994. Base stock level determination for higher cost low demand critical repairable spares. Computers and Operations Research, 21 (4), 411–420.

Ebeling, C.E., 2010. *An introduction to reliability and maintainability engineering.* Long

Grove: Waveland.

Everingham, K., Polaski, G., Riedlin, F., Shirk, M., Deshpande, V., Iyer, A.V., 2008. Operations research enhances supply chain management at the US Coast Guard aircraft repair and supply center. Interfaces, 38 (1), 61-75

Ghobbar, A.A., Friend, C.H., 2002. Sources of intermittent demand for aircraft spare parts within airline operations. Journal of Air Transport Management,8, 221–231.

Ghobbar, A.A., Friend, C.H., 2003. Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model. Computers and Operations Research, 30, 2097–2114.

Ghobbar, A.A., Friend, C.H., 2004. The material requirements planning system for aircraft maintenance and inventory control: a note. Journal of Air Transport Management 10, 217–221.

Ghodrati, B., 2012. Product Support Improvement by Considering System Operating Environment: A Case Study on Spare Parts Procurement. International Journal of Quality and Reliability Management, 29(4), 436-450.

Hadley, G., Whitin, T.M., 1963. Analysis of Inventory Systems. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Kabir, A.B.M.Z., Al-Olayan, A.S., 1994. Joint optimization of age replacement and continuous review spare provisioning policy. International Journal of Operations and Production Management, 14 (7), 53–69.

Kabir, A.B.M.Z., Al-Olayan, A.S., 1996. A stocking policy for spare part provisioning under age based preventive replacement. European Journal of Operational Research, 90 (1), 171–

181.

Kaio, N., Osaki, S., 1978. Optimum ordering policies with lead time for an operating unit in preventive maintenance. IEEE Transactions on Reliability, 27, 270–271.

Kennedy, W.J., Patterson, J.W., Fredendall, L.D., 2001. An overview of recent literature on spare parts inventories. International Journal of Production Economics, 76, 201-215.

Kim, J., Shin, K., Yu, H., 1996. Optimal algorithm to determine the spare inventory level for a repairable-item inventory system. Computers and Operations Research, 23, 289–297.

Kiyak, E., 2012. The effects of aircraft preventive maintenance on reliability. International Journal of Applied Mathematics and Informatics, 6(1), 9-16.

Lau, H.S., Lau, A.H.L., 1995. The multi-product multi-constraint newsboy problem: applications formulation and solution. Journal of Operations Management 13, 153–162.

Lau, H.S., Lau, A.H.L., 1996. The newsstand problem: a capacitated multiple product single period inventory problem. European Journal of Operational Research 94, 29–42.

Ma, L., Zhang, G., 2008. Configuring a supply network in the presence of volume discounts. Journal of Manufacturing Systems, 27, 77-83.

MCTF, 2012. Airline Maintenance Cost: Executive Commentary - An Exclusive Benchmark Analysis (FY2010 data) by IATA's Maintenance Cost Task Force.

Muchiri, A.K. and Smit, K., 2011. Optimizing aircraft line maintenance through task re-clustering and interval de-escalation. Sustainable Research and Innovation Proceedings. Vol 3.

Nahmias, S.,2005. *Production and Operations Analysis*. New York, McGraw-Hill/Irwin.

Natarajan, R., 1968. A reliability problem with spares and multiple repair facilities. Operation Research, 16, 1041–1057.

Regattieri, A., Gamberi,M., Gamberini,R., Manzini,R.,2005. Managing lumpy demand for aircraft spare parts. Journal of Air Transport Management 11, 426-431.

Simpson, V.P., 1978. Optimum solution structure for a repairable inventory problem. Operations Research, 26, 270–281.

Taha, H.A., 2007. *Operations research: an introduction.* New Jersey: Pearson.

Thomas, L.C., Osaki, Shunji, 1978. An optimal ordering policy for a spare unit with lead time. European Journal of Operational Research, 2, 409-419.

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

Wang, W., 2012. A stochastic model for joint spare parts inventory and planned maintenance optimization. European Journal of Operational Research, 216, 127–139.

Zhang, G., 2010. The multi-product newsboy problem with supplier quantity discounts and a budget constraint. European Journal of Operational Research, 206, 350-360.

VITA AUCTORIS

NAME: Jingyao Gu

PLACE OF BIRTH: Shaanxi, China

EDUCATION:

2002-2006 Beihang University, Beijing, China

B.Eng in Aircraft Manufacturing Engineering;

2011-2013 University of Windsor, Windsor, Ontario

M.A.Sc. in Industrial Engineering.

WORK EXPERIENCE:

2007–2011 China Postal Airlines Co., Ltd., Beijing, China

Planning and Purchasing Engineer in Materials and Spare Parts Division