8-2009

Supply Chain Management with Demand Substitution

Laigang Song
University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Laigang Song entitled "Supply Chain Management with Demand Substitution." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering.

Xueping Li, Major Professor

We have read this dissertation and recommend its acceptance:

Alberto Garcia, Rupy Sawhney, Frank Guess

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Laigang Song entitled "Supply Chain Management with Demand Substitution". I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering.

Xueping Li, Co-chair

Alberto Garcia ,Co-chair

We have read this dissertation
and recommend its acceptance:

Rupy Sawhney

Frank Guess

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Supply Chain Management With Demand Substitution

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Laigang Song

August 2009

Dedication

This dissertation is dedicated to my wife and daughter, Dongchun Qiao and Selina Jiayin Song, my parents, Yijun Song and Guiying Han, my brother and sister, Tao Song, Feng Song and Laiwei Song, for always believing in me, inspiring me, and encouraging me.

Acknowledgments

First and foremost, I am deeply indebted to my two advisors Dr. Alberto Garcia and Dr. Xueping Li. Their willingness to support my work and the guidance throughout my studies has allowed me to develop my skills as a researcher within a supportive team environment. I thank them for the precious opportunity.

Within the IIESL Lab, I owe many thanks to my colleagues. Specially, I thank to Dengfeng Yang, Yuerong Chen and Jiao Wang for their supportive team-work. Throughout my work, the administrative staffs, Diana Bishop, Jeannette Myers in the Department of Industrial and Information Engineering also help me greatly.

Abstract

Demand substitution is a very common practice, but due to its inherent difficulty of mathematical modeling, little has been done on the impact of the demand substitution to the supply chain network.This dissertation studied the impact of demand substitution to a supply chain network.

One of the most important measurements of supply chain network, Bullwhip effect, is studied under the demand substitution case. To help understand the influence, a new qualitative measurement of the bullwhip effect is proposed to better capture the essence of the uncertainty associated with supply chain networks. Then a mathematical model is formulated to investigate the bullwhip effect of two products substitution case.

Due to the difficulty of the mathematical modeling of the demand substitution process, “Metamodel” methodology is applied to study the relationship among different aspects of the supply chain network.

Finally, a network based algorithm is proposed to represented the demand substitution process. Graphical Evaluation And Review Technique (GERT) is used to solve the network model. The results demonstrated the effectiveness of the network model to approximate the demand substitution problem. In the end, the dissertation concludes with a summary of the contributions to the state of the art.

Contents

List of Tables

List of Figures

Chapter 1

Literature Review

The literatures of the demand substitution, the bullwhip effect and the meta-model are reviewed in this chapter.

1.1 Demand Process Model

The demand process can be divided into two stages: customer arrival and product selection. The customer arrival is an process that is not affected by product attributes. The product selection process describes the way in which customer selects products.

The two stages can be integrated into one stage to simplify the modeling effort. One widely used demand process model that integrate the two stages together is the one-period autoregressive model *AR(1)*:

$$D_t = d + \rho D_{t-1} + \varepsilon_t \tag{1.1}$$

Where D_t represents the customer demand in period t,d is a positive constant, ρ is the autocorrelation parameter with $|\rho| \leq 1$, and ε_t is the error term that is i.i.d.

(identical and independent distribution) with a symmetric distribution (e.g. normal) having mean zero and variance σ^2 and is uncorrelated with anything known at time $t-1$. The demand process Equation 1.1 was adopted as early as 1987 by Kahn (1987). In recent years, it has been applied to the analysis of the bullwhip effect and information sharing by many researchers. For example Chen et al. (2000a) quantified the bullwhip effect caused by demand forecasting and order lead times in a two-stage supply chain with a manufacturer and a retailer who faces the demand process *AR(1)*. Chen et al. (2000b) considered two demand processes, *AR(1)* and a demand process with linear trend, and they quantitatively analyzed the bullwhip effect for two-stage supply chains consisting of a manufacturer and a retailer. The retailer was assumed to use the exponential smoothing and moving average forecasting techniques to update the mean and standard deviation of demand and thus the retailer's order-up-to point for each period. Leng and Palar (2008) use Equation 1.1 to model the demand process faced by the retailer, compute cost savings generated by information sharing, and conduct a cooperative game analysis for the fair allocation of cost saving in a three-level supply chain. Gilbert (2005) generalized the *AR(1)* process to Autoregressive Integrated Moving Average Model (ARIMA(p,d,q)) and studied the causes of the bullwhip effect.

For two stage demand process model, the two stages need to be modeled separately to study the demand substitution. Poisson process is normally used to model the number of customer arrivals. Normal distribution is used when the variance of customer arrival is not equal to the mean of the customer arrival.

There are two major methods to model the demand substitution process:the rational deterministic allocation model and the dynamic substitution model. The rational deterministic allocation model is widely used in many literatures (McGillivray and Silver, 1978; Parlar and Goyal, 1984; Parlar, 1988; Netessine and Rudi, 2003; Hopp and Xu, 2008; Rajaram and Tang, 2001). In this model, the customer will chose a product based on a predetermined value. Specifically, there is an exogenously random demand for each product. When the demands exceed the inventory, stock-out takes place. In this case, the demand is allocated to other products including leaving without buying based on a predetermined deterministic proportion value.

The second method is not so widely used due to its complexity. In this method, the demand process will be explicitly modeled and studied. The customer is assigned with a utility vector and they will make choices to maximize their utility (Mahajan and van Ryzin, 2001b,a). This method is considerably more complex than the rational deterministic allocation model due to the involvement of a dynamic sample path model.

1.2 Demand Substitution Models

There are two fundamental forms of substitution: "*supplier driven*" and "*consumer driven*". In supplier driven substitution, the decision of product substitution is made by the supplier. Supplier will substitute the product based on its own inventory position, the market forecast and other related information. It often occurs in multi-product manufacturing system and some service industry. In semiconductor industry producing similar integrated circuits with varying performance characteristics; circuits with higher performance characteristics (e.g., speed) could substitute for demand for circuits with lower performance characteristics but not vise versa (Bassok, Yehuda et al., 1999). Another example in the same industry relates to memory chips where a higher capacity (2 Gigabytes) chip can be used to satisfy demands for lower capacity memory (say, 1 Gigabytes). Wagner and Whitin (1958) shows the example from steel industry where steel beams of a higher strength can substitute for beams of lower strength. The supplier can control the inventory by providing promotions, sales and etc in supplier driven substitution. For example, a computer manufacturing finds that 2 Gigabytes memory has too many inventory but 1 Gigabyte's inventory is in critical condition. New shipment will arrive in 3 months without incurring expedite cost. In this case, the supplier can run a promotion to lower the price of 2 Gigabytes memory so that part of the customers who want 1 Gigabyte memory will switch 2 Gigabytes.

In consumer driven substitution, the product substitution decision is made by the consumers. When a product is stock-out, the customer will choose other product or leave without buying. The two methods to model this process have been described in the above section. Due to the heterogeneous customer demands, the supplier is unable to predict the decision of each customer. To prevent product stock-out, the seller must consider all the customer interests to place the inventory replenishment order. In general, consumer driven substitution is more relevant to the inventory policy. Thus this thesis only consider the consumer driven substitution case.

Substitutability is an elegant criterion which deals with the subtle connection between inventory and customer satisfaction. Acceptable levels of inventory depend in part on how substitutable a product is in customers' eyes, that is, how reasonable it is to assume that a customer would cheerfully substitute another available product for the out-of-stock item he or she intended to buy. If an SKU is highly substitutable, inventories can be lowered; inventories for non-substitutable SKUs must be raised. Fuller et al. (1993) discusses the potential advantages of recognizing substitution structures to effectively manage inventories and reduce costs.

The demand substitution is inevitable. The manufacturers try to achieve more market share by expanding their product line and offering more types of products to the customers. To eliminate the huge cost of new product development, the manufacturers will typically change their products over some attributes like the color of shirts, the package size and flavor of the cereal. Having so many choices of similar products, the customers can easily find substitutable items to meet their demand. The substitution can also be caused by the service level agreement. The service level agreement typically involves time limit to a certain piece of equipment. To meet these agreement, substitution is sometime necessary. The laptop computer service industry is good example of substitution. When customer send their laptop computers to the service providers, they will receive an agreement that the laptop will be repaired within some time period. During that period, certain parts might stocked out in the vendor. If they wait for the next replenishment order arrive, some service level agreement can not be maintained. In the long run, it is harmful to violate the service level agreement. The lost of customer goodwill is extremely critical in today highly competitive market. To avoid this violation, substitution using available parts is necessary. These parts will typically have higher specification than that of the original one to maintain customers' goodwill.

Some may argue that if substitution is profitable, it should take place all the way. The reason that substitution is meaningful lay in two aspects. Firstly either no substitution is optimal or complete substitution is optimal. Drezner et al. (1995) proved that the optimal strategies for EOQ model with substitution is partial substitution- a surprising result for most people. Secondly the stock-out can often be observed in practice. Without the policy of substitution, the unsatisfied customers will lead to loss of sales and loss of goodwill. Many literatures have shown that the consideration of demand substitution in the inventory policy can improve the profit as high as 10 percent (Rajaram and Tang, 2001). Considering the enormous investment on the inventory, the improvement can not be ignored by any company.

The existence of the "demand substitution" will finally affect the inventory control policy of the items which are served as possible substitutions for each other. For example, it is not necessary to carry out as much safety sock as if the demand is totally independent since we can rely on the inventory of the substitutable item to satisfy the unmet demand and to prevent from the loss of sales due to the stock out of a particular item. The demand substitution creates interdependency among the items because each demand for currently unavailable item is transferred to its currently in stock substitutes. Optimizing the inventory policy subject to the influence of the interdependency is a very complex problem. But it is clear that the demand substitution has significant influence on the inventory policy (Agrawal and Smith, 2003).

Demand substitution has been a research topic for decades. The first paper study the demand substitution is McGillivray and Silver (1978). On that paper, they investigate the effects of substitutable demand on stocking control policy and the associated inventory/shortage costs. They assumed that all of the substitutable items have the same unit variable cost and shortage penalty. For the parts which are in shorts, the customers will accept another available part with a probability $0 \leq \alpha_{ij} \leq 1$ which is the probability that customer will accept item j as the substitution of item i when item i is out of stock. They investigated the benefits of no substitution and the benefits of completely substitution. The maximum possible savings associated with the selection of the order-up-to level under substitution for each item is obtained. By specially focusing on two-item case, they use simulation and heuristic approach to estimate the order-up-to levels. Their results show that the potential dollar savings in accounting for substitutable demand rise quickly with the inventory level of each products.

Parlar and Goyal (1984) modeled the two-substitutable-product problem as an extension of the single-period problem. For a single-period inventory problem, the result of classical news boy problem (Spearman and Hopp, 2001) can be utilized. Their results showed that the optimal order quantities for each product can be found by maximizing an expected profit function which is strictly concave for a wide range of parameters values.

Drezner et al. (1995) investigated an economic order quantity model with two ordered substitution products. That is, one can be used to substitute the other at a given unit cost. Three cases are studied: no substitution, full substitution and partial substitution. The author argued that the full substitution can not be optimal. Only partial or no substitution may be optimal. By comparing the optimal total cost of these three situations, the author draws the following conclusion.

- the partial substitution is optimal when $c_t \leq (c_{h2} - c_{h1}) \left(\frac{2c_o}{c_{h1}D_1 + c_{h2}D_2}\right)^{1/2}$
- Otherwise “no substitution ” policy will be optimal

Drezner and Gurnani (2000) extended their research (Drezner et al., 1995) from two products to N products. They studied a deterministic nested substitution problem where there are multiple products which can be substituted for each other at a certain cost under an EOQ set-up. They found that when the number of products exceeds two the total cost function may not be convex. A series of variable substitutions lead to an objected function which is proven to be convex and can be optimally solved.

Ingene and Moinzadeh (1993) developed long run profit maximizing stocking and pricing policies in the face of unpredictable but "stationary" demand for a pair of related goods. They examined a profit maximizing company that distributes two related (substitutable) products. The first product is sold in a perfectly competitive market at price p_1. The second product is not held in stock. It is only delivered when one or more customers demand for the item. The second product may be a substitute for the first product for the some customers. The first product resembles the widely distributed products so that the customer knows the fair market price of such item. The second product resembles the high profit product with relatively limited demand so that the fair price of this product is unknown to customers. The goal of their research is to determine the optimal stock level of the first product and the money to charge for the second product to maximize the expected company profit. Since the objective is to maximize the expected company profit, the customers will experience dissatisfaction when the quantity of product or service they received falls short of the level they expected (Parasuraman, A. et al., 1985). This would be unacceptable for some service industries which needs very high Service Level Agreement.

Rajaram and Tang (2001) analyze the impact of product substitution on two key aspects of retail merchandizing: order quantity and expected profits. They extended the news vendor model to include the possibility that a product with surplus inventory can be used to substitute the demand of the out-of-stock products. They use a parameter which range from 0 to 1 to simulate the degree of substitution. They again assume the same holding, shortage, and salvage cost for all products.

Agrawal and Smith (2003) consider the problem of optimizing assortments in a multi-item retail inventory system. The customers will buy items in set. If one item is not available, the customer will either walk away or accept a substitution or change the purchased item set. A demand model to capture this behavior is proposed to derive a tractable approximation of the problem. They assumed a fixed cycle for replenishment with no lead time. This reduces the system to a multi-item news vendor problem, similar to the model mentioned (Ingene and Moinzadeh, 1993). But they consider a larger set of substitutable items.

Bassok, Yehuda et al. (1999) and Bitran, G. R. and Dasu, S. (1992) consider the “one-way substitutability” scenario. They divide the products into several grades. The products in the higher grades can be used to substitute the product in the lower grade with a certain cost. In their model, different holding, shortage and salvage cost across the products are considered. Bitran, G. R. and Dasu, S. (1992) examined the case when product demands are deterministic, but the actual quantity produced is different from the quantity being processed due to random yield. Then they extended their research by letting the product demand be random and adding setup cost for each product substitution. Bassok, Yehuda et al. (1999) shows that a myopic base stock policy is optimal and develop and algorithm for determining and optimal ordering policy.

The similar problem can be found in many other applications. Karaesmen and van Ryzin (2004) studied an overbooking problem with multiple reservation and inventory classes. The inventory classes may be used as a substitution to satisfy the demand of a given reservation class. The object is to maximum the expected profit, given the limit inventory classes. There are two periods in this problem: the reservation period and the service period. During the reservation period, the overbooking level (the number of reservation on hand at the end of reservation period) is decided so that the maximum profit can be obtained. After the reservation period, the cancelation and no-show logic is implemented. Then all the remaining customers are either assigned to an inventory class or are rejected. This is modeled as a transportation problem. Robinson (1995) also considered the time to refuse discount bookings from airline passengers to reserve seats for potential future passengers who are willing to pay a higher fare. The optimal policy is to accept reservation requests as long as the cumulative seats booked do not exceed a given booking limit, when passengers arrive in sequential fare classes. This policy is very similar to the (s, S) inventory model.

Netessine and Rudi (2003) considered two different scenario in the demand substitution: The centralized management where all products are managed by a central decision maker whose objective is to maximize the expected aggregate profit, and decentralized inventory management where each product is managed by and independent decision maker whose objective is to maximizing the expected profit generated by this specific product while interacting with the other decision makers. They proved the objective function with more than two products and full substitution struccture might not be concave and not even quasi-concave. Parlar and Goyal (1984) and Ernst, Ricardo and Kouvelis, Panagiotis (1999) show that the objective function with two or three products is jointly concave.

1.3 Bullwhip Effect

The study of propagation and amplification of the lower level randomness on the higher echelons commonly known as bullwhip effect in a supply chain has been

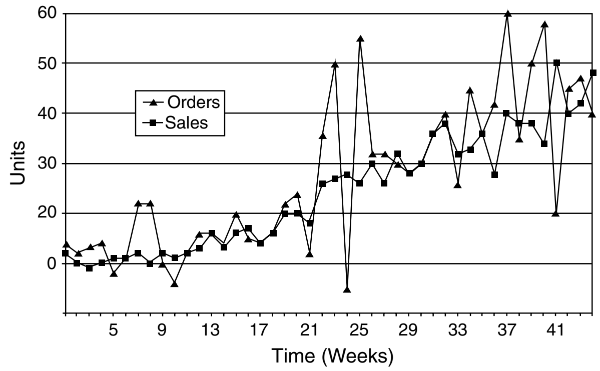


Figure 1.1: The distortion of demand information

of interest from viewpoints of both design and operation of the chain. This variance of demand amplification was first proposed and studied by Forrester (Forrester, 1958). This paper laid the foundation stone for the following researches. Many authors are inspired from observation to develop management games to demonstrate the variance amplification. The most successful game is the well-known Beer Game (Sterman, 1989).

A supply chain typically consists of manufacturers, distributors and final retailers. Only retailers have direct access to final customers. As to distributors, only orders from retailers can be seen. Real consumers' demands are not visible to distributors. The same phenomenon happens at upper levels of the supply chain. When products are distributed downward along the supply chain, information flows upward towards networks. It has been noticed that the information will be distorted due to various reasons. This distortion of the demand in the upstream of the supply chain is widely known as "bullwhip" effect (Chen et al., 2000b). Figure 1.1 shows the distortion of demand information observed in the retailer. The bullwhip effect has been commonly and conveniently measured as ratio of variances of lead time demand to that of end customer demand. It depends on the demand process at the end customer level, lead time demand, forecasting models used, and replenishment policy employed at various levels. Lee et al. (1997a,b)give five important causes for the bullwhip effect: the use of 'demand signal processing', non-zero lead times, order batch, supply shortages and price fluctuations.

The distortion of demand information implies that the entity in the supply chain who only observes its immediate order data will be misled by the amplified demand patterns. This distorted information will lead to inefficiencies in many parts of the supply chain, such as excess raw materials due to unplanned purchases from suppliers, additional manufacturing expenses created by excess order demands, inefficient utilization and overtime, excessive warehouse cost and so on. Fuller et al. (1993) pointed out that inefficiencies bear part responsibility for $75 billion to $100 billion worth of inventory caught between various members of the $300 billion grocery industry in 1993.

(Lee et al., 2004) gives the quantitative analysis on the bull whip effect under the order-up-to-S policy. Metters (1997) explores the magnitude of the bullwhip effect by establishing an empirical lower bound on the profitability under different level of demand variance. They model the supply chain as a periodic, time-varying, stochastic demand dynamic program with capacitated production. Dynamic programming is used to determine the optimal ordering policy. Chen et al. (2000b) investigate the bullwhip effect on a two-stage supply chain consisting of only one retailer and one manufacturer. The demand forecasting and order lead times are considered in their model. Song et al. (2008) investigate the bullwhip effect under demand substitution.

1.4 Metamodel

Barton (1992) defines a metamodel as a model of the simulation model to expose the fundamental relationship between the input and output. Following this definition, many famous theorems can be viewed as metamodels. For example, the Little's Law $L = \lambda W$ is a metamodel. Santos and Santos (2007) define a metamodel as abstractions of a simulation model that exposes the system's input-output relationship through simple mathematical expressions.

The metamodel can be constructed from the collected data using different methods. The linear polynomial approximations are the most widely used methods due to its simplicity (Kleijnen, 2007a,b). Nonlinear techniques are more complex but more flexible and powerful. Nonlinear regression is the most studied nonlinear techniques (Reis dos Santos and Porta Nova, 2006). Other methods include (Kleijnen, 2007a), neural network (Badiru and Sieger, 1998), Bayesian approaches (Cheng, 1999).

Metamodels can be built through the following steps. 1)Understand the problem. 2) Decide input values. The input values should be designed based on the principals of the design of experiments. A robust design experiment works well across a broad range of scenarios and provides solutions that are less likely to produce unexpected results (Kleijnen and van Beers, 2005). The detailed input design method can be found (Kleijnen, 2008). 3)Build the simulation model. In most cases, the simulation model means computer based simulation. 4) Collect the response data. The resulting simulation responses are collected from the experiments. 5) Estimate the parameters. The appropriate mathematical model with few parameters will be calibrated so that the responses can be fitted.

With the increasing power and decreasing cost of the computer, computer simulation has the ability to run a large amount of experiments. This is very different from the real world experiments. In real experiments, the number of factors or the possible values of each factor are limited due to the high cost of experiment. However, these constraints do not apply for computer simulation. Kleijnen et al. (2005) gives a detailed discussion on the similarities and the differences between computer simulation and the real experiments.

Metamodels are very useful in many aspects. Firstly, they can show the basic properties of the target system. Findings from a metamodel can be used to verify and validate the model. Secondly, a metamodel can be used to identify the factors which has the most significant impact of the output. Thirdly, due to the low computer resource requirement, a metamodel can be replicated for many times to test different scenarios. This is particularly important when the output is random. That is, when the inputs of a simulation are constant, the effect of inputs on outputs can not be estimated. When the original model is a component of a very complex system, metamodel can be used to substitute the original component. In some case, it is the only solution.

Metamodels are very useful in many aspects. Firstly, they can show the basic properties of the target system. Findings from a metamodel can be used to verify and validate the model. Secondly, a metamodel can be used to identify the factors which has the most significant impact of the output. Thirdly, due to the low computer resource requirement, a metamodel can be replicated for many times to test different scenarios. This is particularly important when the output is random. That is, when the inputs of a simulation are constant, the effect of inputs on outputs can not be estimated. When the original model is a component of a very complex system, metamodel can be used to substitute the original component. In some case, it is the only solution.

The fierce competition in the market forces all practitioners to pay attention to not only themselves but also all entities in the supply chain. The ultimate goal is to maintain customer relationship at a high level. Customers always expect their demand to be satisfied as soon as possible. Delay in customer demand will reduce the customer's goodwill and finally lose customer. Good relationships between suppliers and customers are built over long time cooperation. There are many criteria to measure the level of a service. Among all the criteria, service level agreement (SLA) is the most commonly used one. The SLAs are different in different areas and different commodities, but they typically involve the service time limit related to a certain piece of equipment or service. The so-called $80/20$ rule, meaning 80% of all customers must be served within 20 seconds, is very popular in call center industry (Milner and Olsen, 2008). The actual number may vary but the form is about the same (Application Service Provider Industry Consortium, 2000). For large companies, the vast amount of customers leads to various service levels. For small companies, the situation may not be so complex, but still the companies need to be prepared for multiple service level requirements.

To provide better customer service experience, the simplest way is to increase the on-hand inventory for any potential demand. For manufacturing industries, producing more products will increase the possibility of satisfying the customers' demand instantly. For service industries, maintaining more spare parts will result in shorter service time. But on-hand inventory level can not be too high due to the inevitable holding cost and purchasing cost. For some expensive items, having a high level of inventory means that a high amount of capital is locked by these items (Hopp and Spearman, 2000). Thus, the annual revenue from the capital is reduced. This contradiction is faced by almost every manager in the world. Specifically, the decisions faced by supply managers in procuring and positioning parts to satisfy these complex service agreements at minimum inventory investment have become particularly difficult (Caggiano et al., 2007).

Much work has been done in the inventory optimization of a multi-echelon system under the constraints of the time based service level agreements. Caggiano et al. (2007) investigates a continuous review inventory model for a multi-item multi-echelon service parts distribution system in which a time-based service level requirement exists. By driving exact time-based fill-rate expressions for each item within its distribution channel, an intelligent greedy algorithm for large-scale problems is proposed. The simulation results show that their algorithm is efficient to large-scale problems. Caglar et al. (2004) examines a multi-item, two-echelon system with the objective of minimizing total system inventory investment subjected to an average response time constraint. They propose a method based on Lagrangian decomposition. The experiments show the method is efficient for fairly large problems. Wong et al. (2007) investigate the same problem as in (Caglar et al., 2004). The experiments show that their greedy algorithm has better computation performance. With the help of some heuristics, their algorithm can be even more efficient. Wong et al. (2008) proposed new measurement of supply chain efficiency. Craven and Islam (2007) surveyed the recent development of applying operation research in supply chain management.

Chapter 2

Overall Conceptual Approach

The overall conceptual approach for proposed methodology is summarized in Figure 2.1. The overall conceptual approach can be divided into 3 phases.

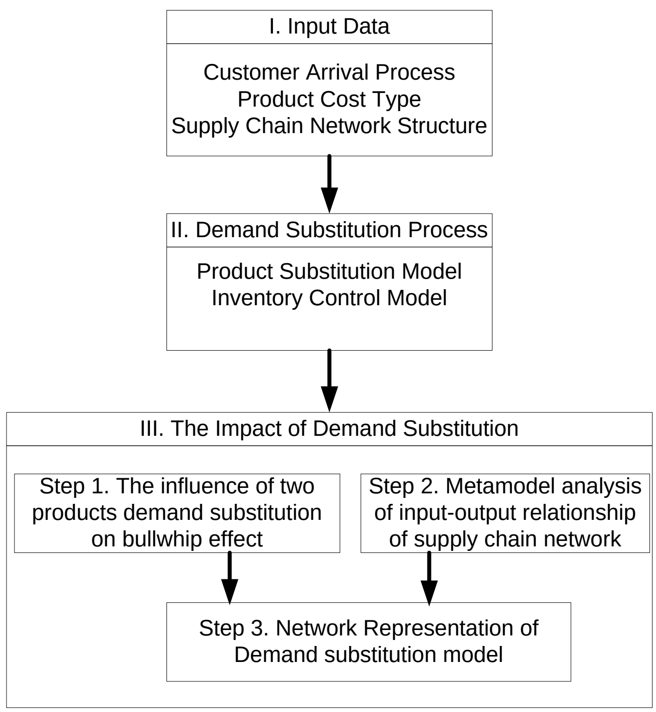


Figure 2.1: Overall Conceptual Approach

Phase I Input Data The required inputs are: Customer arrival process (Poisson or Normal); Product cost types(unit cost, carrying cost and salvage cost); Supply chain network structure (single echelon, multi-echelon).

Phase II Demand Process Model In this phase, two important decisions need to be made: the product substitution model(Rational deterministic allocation model or Dynamic substitution model); the inventory control model (EOQ, (s,S), order-up-to and etc.).

Phase III The impact of demand substitution This phase is composed of 3 steps:

Step 1 The influence of two products demand substitution on bullwhip effect. The dissertation starts from the simple case with only two products. One products is used to substitute the other product. The lower limit of the bullwhip effect is obtained in this case.

Step 2 Metamodel analysis of input-output relationship of supply chain network. Due to the difficulty of mathematical formulation of demand substitution process, metamodel analysis is utilized to study the input output relationship of a three echelon supply chain network.

Step 3 Network Representation of Demand substitution model. By replacing the stochastic inventory changing process with deterministic service rate, a network representation of the demand process is established for l products.

Chapter 3

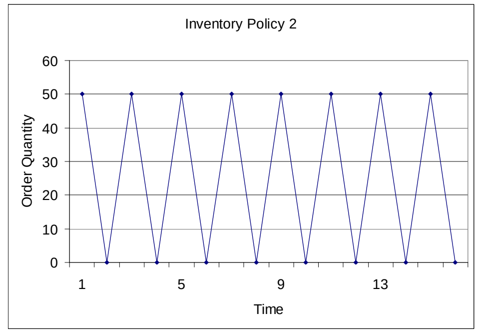
The Impact of Demand Substitution to Bullwhip Effect

3.1 The Measurement of Bullwhip Effect

3.1.1 Problem Definition

The bullwhip effect has been commonly and conveniently measured as ratio of variances of lead time demand to that of end customer demand. However, this definition is not comprehensive. Kawagoe and Wada (2005) reports an example that the traditional variance bullwhip measurement fails. They then propose descriptive statistics to measure the bullwhip effect. Furthermore, they take the frequency of the variability into consideration by using the concept of stochastic dominance. All these measurements are based on statistical methodologies. Figure 3.1 shows the order quantity patterns by applying two different inventory polices. The variances of two order quantity are the same, but clearly the second policy (Figure 3.1(b)) incurs more order quantity changes than the first inventory (Figure 3.1(a)) policy does. If high costs are related to the changes in

(a) Inventory Policy 1



(b) Inventory Policy 2

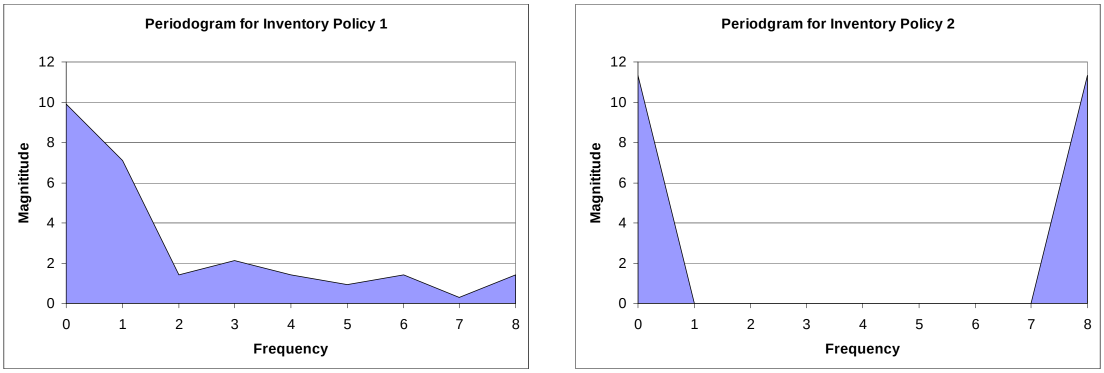
Figure 3.1: The order quantity patterns of two different inventory policies

the order quantity, then the second inventory policy will cost more than the first policy. The traditional variance based bullwhip effect measurement fails since the frequency of the changes is not considered. In reality, the order quantity patterns like Figure 3.1 are very unlikely to happen, but similar situations will be encountered.

A supply chain consists of many parts with complex interactions among them. It can be seen as a system, thus the methodologies of control engineering are applicable in the analysis of supply chains. The control system engineering approach is one of the two approaches to tackle the inventory replenishment problems. The other approach is the traditional statistical inventory control approach. In this section, we propose a new quantitative measurement of the bullwhip effect based on the control engineering approach.

Spectral Analysis

Spectral analysis is a mathematical technique used to decompose a time series into constituent frequencies or periodicities. The amplitude or variance associated with each frequency component is known as the 'spectral density estimate'. The result is a plot of the amplitude versus the frequencies which is called the



(a) Periodogram of Inventory Policy 1 (b) Periodogram of Inventory Policy 2

Figure 3.2: Periodogram of two different order quantity patterns in Figure 3.2

'periodogram'. Figure 3.2 shows the Periodogram of two difference order quantity patterns mentioned in Figure 3.1. It is clearly that these two order patterns are totally different in Periodogram. Although various methods are available to conduct spectral analysis, a technique called Fast Fourier Transformation is often used.

Fast Fourier Transformation (FFT)

The Fourier transform is an algebraic method of decomposing any time series into a set of pure sinuous waves of different frequencies, with a particular amplitude and phase angle associated with each frequency. The algebraic sum of the sinusoidal components, adjusted for phase angle can accurately reproduce the original time series. Suppose that a physical process is represented by a function of time, $h(t)$. This function is sampled at N times, $t_k = k\delta t$, where $k = 0, 1, 2, \cdots, N-1$. From these N measurements, h_k, N complex amplitudes, H_n, are determined which satisfy the N equations:

$$H_n = \sum_{k=0}^{N-1} h_k e^{ik\frac{2\pi n}{N}} \tag{3.1}$$

The Fourier analysis routine calculates the complex coefficients H_n from the time series data h_k. These concepts can be applied in inventory management area by considering the demand data as samples of an underlying physical demand process. Many distinct FFT algorithms involve a wide range of mathematics, from simple complex-number arithmetic to group theory and number theory. A wide variety of software packages are available on the Internet. Microsoft Excel provides it own Fourier analysis function via Analysis ToolPak add-in* . We use this package to analyze the demand processes and the order quantity patterns.

The New Measurement: Inventory Entropy

Both "Variance" and "Entropy" are the measurement of the disorder and uncertainty associated with a system. Entropy has been widely used in many areas such as information theory, thermodynamics, and astrophysics and so on. We borrow the term "Entropy" here to describe the "Bullwhip" effect in supply chains. We borrow this term to measure the degree of uncertainty before and after the inventory control policy. Just like any other signal, the demand information is changed by different inventory control policy. The corresponding change can be measured in different ways. Our measurement is different from others due to the consideration in time and frequency domain.

We define the inventory entropy as the area under the curve in "Periodogram"

$$I = \sum_{n=-\infty}^{\infty} A_n \omega_n \tag{3.2}$$

Where A is the amplitude and ω is the frequency. The ratio of the entropy of the demand process to order quantity is an new quantitative measurement of the bullwhip effect.

*http://office.microsoft.com/en-us/excel/HP052038731033.aspx

3.1.2 Numerical Experiment

The new bullwhip measurement will be applied to a supply chain with one retailer and one manufacturer. The retailer will use moving average algorithm to forecast future demands. Suppose the order cost is negligible, thus the Order-Up-To policy is adopted so that the result can be compared with results from other literatures. We use the same demand process as Chen et al. (2000b) which is widely used in many literatures.

$$D_t = d + \rho D_{t-1} + \varepsilon_t$$

Where μ is the average demand, ρ is the correlation parameter with $|\rho| < 1$, and the error term, ε_t are independent and identically distributed from a normal distribution with mean 0 and variance σ^2 . The order up to level is decided as:

$$y_t = \hat{D}_t^L + z\hat{\sigma}_{et}^L$$

where $\hat{D}_t^L$ is an estimate of the mean lead time demand, $\hat{\sigma}_{et}^L$ is an estimate of the standard deviation of the L period forecast error, and z is a constant chosen to meet a desired service level. We use simple moving average to estimate $\hat{D}_t^L$ and $\hat{\sigma}_{et}^L$ based on the demand observations from previous p periods.

$$\hat{\sigma}_{et}^L = L\left(\frac{\sum_{i=1}^{p} D_{t-i}}{p}\right) \qquad \hat{\sigma}_{et}^L = C_(L,p)\sqrt{\frac{\sum_{i=1}^{p}(e_{t-i})^2}{p}}$$

where $e_t = D_t - \hat{D}_t^l$ and $C_{I,p}$ is a constant function of L, ρ and p (Ryan, 1997). The order quantity is

$$q_t = y_t - Y_{t-1} + D_{t-1}$$

and we make the assumption that the excess inventory is returned without cost.

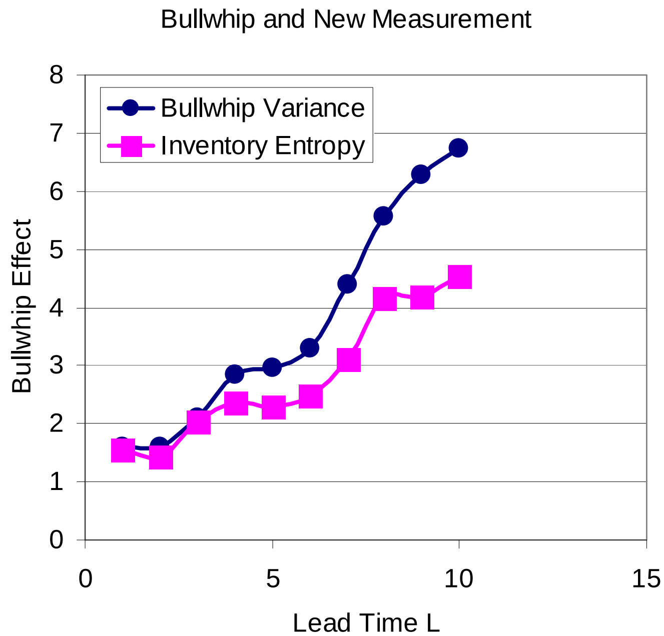
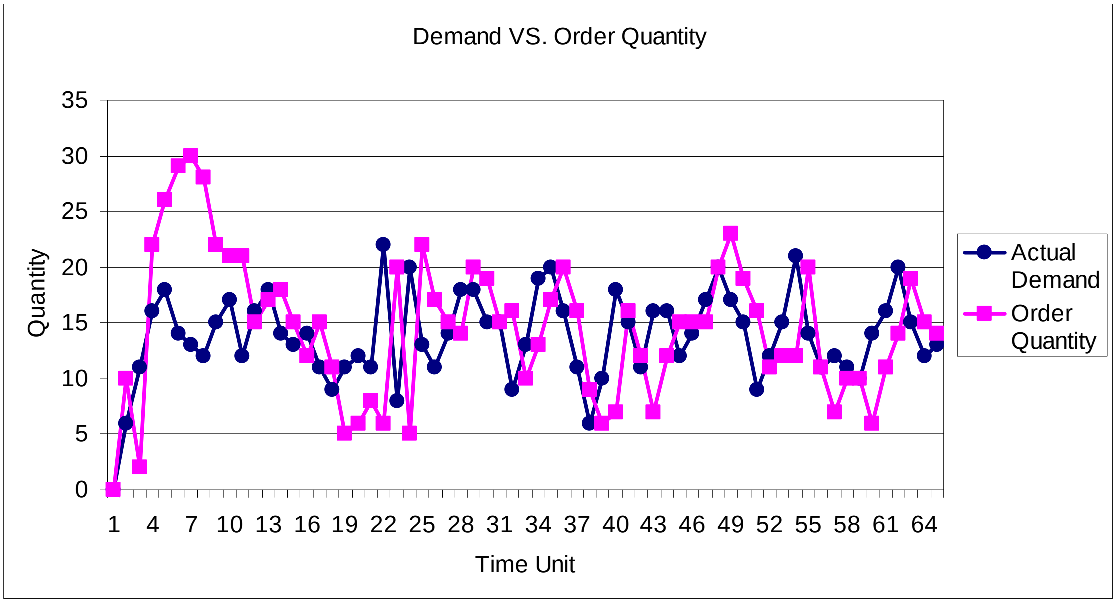


Figure 3.3: The impact of lead time on bullwhip effect.

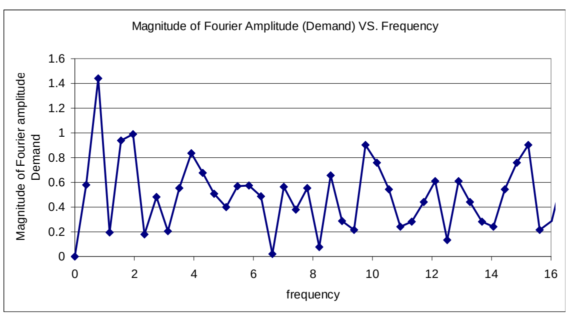
The model is built on a spreadsheet. We arbitrarily set the value of the above parameters as follows: $\mu = 10, \rho = 0.25, \mu_E = 0, \sigma_E = 3, p = 5, L = 3, z = 1.64$, and $C_L, p = 1.2$. The demand estimate for period 1 is set to 10, the average demand per period. FFT technology is used to draw the periodogram.

Figure 3.3 shows the relationship between the lead time and the bullwhip effect measured both by variance and the inventory entropy. The bullwhip effect is proportional to the lead time. The more lead time, the more significant bullwhip effect is. (Conduct a comparison and do a discussion on the difference of the results by the new measurement and by the previous method.)

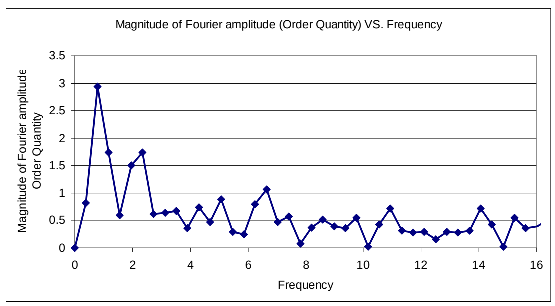
Figure 3.4 shows the time series data of the demand process and the corresponding order quantity. The periodogram of both time series data are shown as well. The periodogram of the demand process shows that the amplitude is distributed more evenly on all frequencies and the periodogram shows left skewed distribution among all frequencies. This is coincided with the demand process



(a) Actual Demand VS. Order Quantity



(b) Fourier Amplitude (Demand) VS. Frequency



(c) Fourier amplitude (Order Quantity) VS. Frequency

Figure 3.4: Demand process, order quantity process and the periodograms

we used in the experiments since normally distributed time series data leads to more evenly distributed amplitude on periodogram.

3.2 The Influence of Demand Substitution to a Two-Echelon Supply Chain

A simple supply chain in which in each period t, a single retailer observes the inventory level for two products 1, 2 and places orders q_{ti}, $i = 1, 2$, to a single distributor. After the order is placed the retailers observes and fills customer demand for that period, denoted by D_{ti}, $i = 1, 2$. Product 1 can be used to substitute product 2. In this model, we assume that a fixed percentage of product 1 is used to substitute product 2. This percentage is denoted as λ,$0 \leq \lambda < 1$. The lead time for both products are the same, L. The order placed at the end of period t is received at the start of period $t + L$. The customer demands seen by the retailer (the lowest echelon of the supply chain) are random variables of the form:

$$\begin{aligned} D_{t1} &= \mu_1 + \rho_1 D_{t-1,1} + \epsilon_{t1} + \lambda D_{t1} \\ D_{t2} &= \mu_2 + \rho_2 D_{t-1,2} + \epsilon_{t2} - \lambda D_{t1} \end{aligned} \tag{3.3}$$

where μ_i is a non-negative constant, ρ_i is a correlation parameter with $|\rho_i| < 1$ and ϵ_{ti} are independent and identically distributed(i.i.d.) from a symmetric distribution with mean 0 and σ_i^2, $i = 1, 2$. From eqn. 3.3, the following results can be obtained.

$$
\begin{aligned}
E(D_{t1}) &= \frac{\mu_1}{1-\rho_1-\lambda} && \text{(3.4a)}\\
Var(D_{t1}) &= \frac{\sigma_1^2}{(1-\lambda)^2-\rho_1^2} && \text{(3.4b)}\\
E(D_{t2}) &= \frac{\mu_2(1-\rho_1-\lambda)-\lambda\mu_1}{(1-\rho_1-\lambda)(1-\rho_2)} && \text{(3.4c)}\\
Var(D_{t2}) &= \frac{[(1-\lambda)^2-\rho_1^2]\sigma_2^2-\lambda^2\sigma_1^2}{[(1-\lambda)^2-\rho_1^2](1-\rho_2^2)} && \text{(3.4d)}
\end{aligned}
$$

Assume the retailer follows a simple *Order-up-to* inventory policy, the order-up-to point is estimated as:

$$
\begin{aligned}
y_{t1} &= D_{t1}^L + z_1\hat{\sigma}_{et_1}^L\\
y_{t2} &= D_{t2}^L + z_2\hat{\sigma}_{et_2}^L
\end{aligned} \tag{3.5}
$$

where $D_{ti}^L,\ i=1,2$ is an estimation of the mean lead time demand, $\sigma_{et_i}=c\sigma_i^L$, $i=1,2$ is an estimate of the standard deviation of the L period forecast error, and z_i is a constant chosen to meet a desired service level. Note that $\sigma_e^L=c\sigma^L$ for some constant $c>1$.

Suppose the retailer uses a simple moving average to estimate D_t^L and σ_{et}^L based on the information of the past p periods. Then

$$
\hat{D}_t^L = L\left(\frac{\sum_{i=1}^p D_{t-i}}{p}\right) \tag{3.6}
$$

$$
\hat{\sigma_{et}^L} = C_{L,\rho}\sqrt{\frac{\sum_{i=1}^p (D_{t-i}-\hat{D}_{t-i})^2}{p}} \tag{3.7}
$$

where $C_{L,p}$ is a constant function of L and ρ, $D_{t-i}-\hat{D}_{t-i}$ is the forecast error of the $(t-i)^{th}$ period. We assume that both product have the same predication periods. That is p is the same to both products.

We are interested in the impact of demand substitution to the Bullwhip effect. If we can get the relationship between the orders to the manufacturer on time period t and the demand of time period t or $t-1$, then we are able to find the variance of the orders placed by the lowest echelon station to the up-level station.

As we already know the order-up-to inventory point is y_t. Thus, at the beginning of each period, the object inventory level is known. To replenish the inventory level to y_t, an order with quantity $y_t - (y_{t-1} - D_{t-1})$ is needed. $y_{t-1} - D_{t-1}$ is the amount of inventory left from the last period. The substitutions are also taken into consideration. Thus, the demand for the two products are as follows.

$$q_{t,1} = y_{t,1} - (y_{t-1,1} - D_{t-1,1}) = y_{t,1} - y_{t-1,1} + D_{t-1,1} \qquad \text{Product 1}$$

$$q_{t,2} = y_{t,2} - (y_{t-1,2} - D_{t-1,2}) = y_{t,2} - y_{t-1,2} + D_{t-1,2} \qquad \text{Product 2}$$

Note that q_t might be negative if the remained inventory from last period is greater thatn the order-up-to point of this period. In this case, we assume that this is the excess inventory and can be returned without any cost. Thus no inventory holding cost is considered to simplify our model. Chen et al. (2000a) proved that *Var(q)* and *Var(q+)* are quite close. Here $q^{+} = max\{q_t, 0\}$.

Given the equation of the order quantity, demand forecast and the standard deviation of the L period forecast error, we can further express q_t as follows:

$$
\begin{aligned}
q_{t,1} &= y_{t,1} - y_{t-1,1} + D_{t-1,1} \\
&= (\hat{D}^L_{t,1} + z_1\hat{\sigma}^L_{et,1}) - (\hat{D}^L_{t-1,1} + z_1\hat{\sigma}^L_{et-1,1}) + D_{t-1,1} \\
&= (\hat{D}^L_{t,1} - \hat{D}^L_{t-1,1}) + D_{t-1,1} + z_1(\hat{\sigma}^L_{et,1} - \hat{\sigma}^L_{et-1,1}) \\
&= \frac{L}{p}(\sum_{i=1}^{p} D_{t-i,1} - \sum_{i=1}^{p} D_{t-1-i,1}) + D_{t-1,1} + z_1(\hat{\sigma}^L_{et,1} - \hat{\sigma}^L_{et-1,1}) \\
&= \frac{L}{p}(D_{t-1,1} - D_{t-p-1,1}) + D_{t-1,1} + z_1(\hat{\sigma}^L_{et,1} - \hat{\sigma}^L_{et-1,1}) \\
&= (1 + L/p)D_{t-1,1} - (L/p)D_{t-p-1,1} + z_1(\hat{\sigma}^L_{et,1} - \hat{\sigma}^L_{et-1,1})
\end{aligned}
$$

Then the variance of the order quantity $q_{t,1}$ for product 1 at time period t is:

$$
\begin{aligned}
Var(q_{t,1}) &= Var[(1 + L/p)D_{t-1,1} - (L/p)D_{t-p-1,1} + z_1(\hat{\sigma}^L_{et,1} - \hat{\sigma}^L_{et-1,1})] \\
&= (1 + L/p)^2 Var(D_{t-1,1}) - 2(L/p)(1 + L/p) \times Cov(D_{t-1,1}, D_{t-p-1,1}) \\
&\ + (L/p)^2 Var(D_{t-p-1,1}) + z_1^2 Var(\hat{\sigma}^L_{et,1} - \hat{\sigma}^L_{et-1,1}) \\
&\ + 2z_1(1 + 2L/p) \times Cov(D_{t-1,1}, \hat{\sigma}^L_{et,1}) \\
&= \left(1 + 2\frac{L}{p} + 2\frac{L^2}{p^2}\right) Var(D_1) - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) Cov(D_{t-1,1}, D_{t-p-1,1}) \\
&+ z_1^2 Var(\hat{\sigma}^L_{et,1} - \hat{\sigma}^L_{et-1,1}) + 2z_1(1 + 2L/p) Cov(D_{t-1,1}, \hat{\sigma}^L_{et,1})
\end{aligned}
\tag{3.8}
$$

To further simply Equation 3.8, we need to calculate $Cov(D_{t-1,1}, D_{t-p-1,1})$ and $Cov(D_{t-1,1}, \hat{\sigma}^L_{et,1})$

$$\begin{aligned} Cov(D_{t-1,1}, D_{t-p-1,1}) &= Cov(\frac{1}{1-\lambda}(\mu_1 + \rho_1 D_{t-2,1} + \epsilon_{t,1}), D_{t-p-1,1}) \\ &= \underbrace{Cov(\frac{\mu_1}{1-\lambda}, D_{t-p-1,1})}_{=0} + \frac{\rho_1}{1-\lambda} Cov(D_{t-2,1}, D_{t-p-1,1}) \\ &+ \frac{1}{1-\lambda} \underbrace{Cov(\epsilon_{t,1}, D_{t-p-1,1})}_{=0} \\ &= \frac{\rho_1}{1-\lambda} Cov(D_{t-2,1}, D_{t-p-1,1}) \\ &\vdots \\ &= \left(\frac{\rho_1^p}{(1-\lambda)^p}\right) Cov(D_{t-p-1,1}, D_{t-p-1,1}) \\ &= \frac{\rho_1^p}{(1-\lambda)^p} Var(D_1) \end{aligned} \tag{3.9}$$

Note that $Cov(\frac{\mu_1}{1-\lambda}, D_{t-p-1,1}) = 0$ because $\frac{\mu_1}{1-\lambda}$ is a constant, $Cov(X, a) = 0$. $Cov(\epsilon_{t,1}, D_{t-p-1,1}) = 0$ because $\epsilon_{t,1}, D_{t-p-1,1}$ are independent from each other.

Ryan (1997) proved the following result which can further simplify Equation 3.8.Assume the customer demands seen by a retailer are random variables of the form given like $D_{t1} = \mu_1 + \rho_1 D_{t-1,1} + \epsilon_{t1}$ and the error terms ϵ_t are i.i.d. from a symmetric distribution with mean 0 and variance σ^2. Let the estimate of the standard deviation of the L period forecast error be $\sigma^L_{et} = C_{L,\rho}\sqrt{\frac{\sum_{i=1}^{p}(D_{t-i} - \hat{D}_{t-i})^2}{p}}$. Then

$$Cov(D_{t-i}, \hat{\sigma}^L_{et}) = 0, \forall i = 1, \ldots, p \tag{3.10}$$

By apply the results of Equation 3.9 and Equation 3.10, the expression about $Var(q_{t,1})$ can be further simplified. Since two products have different variance and expectation, they will be discussed separately.

$$
\begin{aligned}
Var(q_{t,1}) &= \left(1 + 2\frac{L}{p} + 2\frac{L^2}{p^2}\right) Var(D_1) - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) Cov(D_{t-1,1}, D_{t-p-1,1}) \\
&+ z_1^2 Var(\hat{\sigma}_{et,1}^L - \hat{\sigma}_{et-1,1}^L) + 2z_1(1 + 2L/p) Cov(D_{t-1,1}, \hat{\sigma}_{et,1}^L) \\
&= \left(1 + 2\frac{L}{p} + 2\frac{L^2}{p^2}\right) Var(D_1) - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) \frac{\rho_1^p}{(1-\lambda)^p} Var(D_1) \\
&+ z_1^2 Var(\hat{\sigma}_{et,1}^L - \hat{\sigma}_{et-1,1}^L) + 0 \\
&= Var(D_1) \left[1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)\left(1 - \frac{\rho_1^p}{(1-\lambda)^p}\right)\right] + z_1^2 Var(\hat{\sigma}_{et,1}^L - \hat{\sigma}_{et-1,1}^L)
\end{aligned} \tag{3.11}
$$

For product 2, we apply the same procedures as product 1.

$$
\begin{aligned}
q_{t,2} &= y_{t,2} - y_{t-1,2} + D_{t-1,2} \\
&= (\hat{D}_{t,2}^L + z_1 \hat{\sigma}_{et,2}^L) - (\hat{D}_{t-1,2}^L + z_1 \hat{\sigma}_{et-1,2}^L) + D_{t-1,2} \\
&= (\hat{D}_{t,2}^L - \hat{D}_{t-1,2}^L) + D_{t-1,2} + z_2(\hat{\sigma}_{et,2}^L - \hat{\sigma}_{et-1,2}^L) \\
&= \frac{L}{p}\left(\sum_{i=1}^{p} D_{t-i,2} - \sum_{i=1}^{p} D_{t-1-i,2}\right) + D_{t-1,2} + z_2(\hat{\sigma}_{et,2}^L - \hat{\sigma}_{et-1,2}^L) \\
&= \frac{L}{p}(D_{t-1,2} - D_{t-p-1,2}) + D_{t-1,2} + z_2(\hat{\sigma}_{et,2}^L - \hat{\sigma}_{et-1,2}^L) \\
&= (1 + L/p) D_{t-1,2} - (L/p) D_{t-p-1,2} + z_2(\hat{\sigma}_{et,2}^L - \hat{\sigma}_{et-1,2}^L)
\end{aligned}
$$

$$
\begin{aligned}
Var(q_{t,2}) &= Var[(1+L/p)D_{t-1,2} - (L/p)D_{t-p-1,2} + z_2(\hat{\sigma}^L_{et,2} - \hat{\sigma}^L_{et-1,2})] \\
&= (1+L/p)^2 Var(D_{t-1,2}) - 2(L/p)(1+L/p) \times Cov(D_{t-1,2}, D_{t-p-1,2}) \\
&\quad + (L/p)^2 Var(D_{t-p-1,2}) + z_2^2 Var(\hat{\sigma}^L_{et,2} - \hat{\sigma}^L_{et-1,2}) \\
&\quad + 2z_2(1+2L/p) \times Cov(D_{t-1,2}, \hat{\sigma}^L_{et,2}) \\
&= \left(1 + 2\frac{L}{p} + 2\frac{L^2}{p^2}\right) Var(D_2) - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) Cov(D_{t-1,2}, D_{t-p-1,2}) \\
&\quad + z_2^2 Var(\hat{\sigma}^L_{et,2} - \hat{\sigma}^L_{et-1,2}) + \underbrace{2z_2(1+2L/p)Cov(D_{t-1,2}, \hat{\sigma}^L_{et,2})}_{=0}
\end{aligned}
\tag{3.12}
$$

Now we calculate $Cov(D_{t-1,2}, D_{t-p-1,2})$

$$
\begin{aligned}
Cov(D_{t-1,2}, D_{t-p-1,2}) &= Cov(\mu_2 + \rho_2 D_{t-2,2} + \epsilon_{t-1,2} - \lambda D_{t-1,1}, D_{t-p-1,2}) \\
&= \underbrace{Cov(\mu_2, D_{t-p-1,2})}_{=0} + \rho_2 Cov(D_{t-2,2}, D_{t-p-1,2}) \\
&\quad + \underbrace{Cov(\epsilon_{t-1,2}, D_{t-p-1,2})}_{=0} - \lambda Cov(D_{t-1,1}, D_{t-p-1,2}) \\
&= \rho_2 Cov(D_{t-2,2}, D_{t-p-1,2}) - \lambda Cov(D_{t-1,1}, D_{t-p-1,2}) \\
&\vdots \\
&= \rho_2^p Var(D_2) - \lambda \sum_{i=0}^{p-1} \rho_2^i Cov(D_{t-1-i,1}, D_{t-p-1,2})
\end{aligned}
$$

We assume that the covariance is only affected by the number of periods which are taken into consideration. That is

$$
\begin{aligned}
Cov(D_{t,1}, D_{t-p,2}) &= Cov(D_{t-1,1}, D_{t-p-1,2}) \\
D_{t,1} &= \frac{1}{1-\lambda}(\mu_1 + \rho_1 D_{t-1,1} + \epsilon_{t,1}) \\
D_{t-p,2} &= \mu_2 + \rho_2 D_{t-p-1,2} + \epsilon_{t-p,2} - \lambda D_{t-p,1} \\
Cov(D_{t,1}, D_{t-p,2}) &= Cov(\frac{1}{1-\lambda}(\mu_1 + \rho_1 D_{t-1,1} + \epsilon_{t,1}), \mu_2 + \rho_2 D_{t-p-1,2} + \epsilon_{t-p,2} - \lambda D_{t-p,1}) \\
&= \frac{\rho_1\rho_2}{1-\lambda} Cov(D_{t-1,1}, D_{t-p-1,2}) - \frac{\lambda\rho_1}{1-\lambda} Cov(D_{t-1,1}, D_{t-p,1}) \\
Cov(D_{t-1,1}, D_{t-p-1,2}) &= \frac{\rho_1\rho_2}{1-\lambda} Cov(D_{t-1,1}, D_{t-p-1,2}) - \frac{\lambda\rho_1}{1-\lambda} Cov(D_{t-1,1}, D_{t-p,1}) \\
Cov(D_{t-1,1}, D_{t-p-1,2}) &= \frac{-\lambda(1-\lambda)\rho_1^p}{(1-\lambda-\rho_1\rho_2)(1-\lambda)^p} Var(D_1) \\
Cov(D_{t-2,1}, D_{t-p-1,2}) &= \frac{-\lambda(1-\lambda)\rho_1^{p-1}}{(1-\lambda-\rho_1\rho_2)(1-\lambda)^{p-1}} Var(D_1) \\
&\vdots \\
Cov(D_{t-i,1}, D_{t-p-1,2}) &= \frac{-\lambda(1-\lambda)\rho_1^{p-i+1}}{(1-\lambda-\rho_1\rho_2)(1-\lambda)^{p-i+1}} Var(D_1) \\
&\vdots \\
Cov(D_{t-i,1}, D_{t-p-1,2}) &= \frac{-\lambda(1-\lambda)\rho_1^{2}}{(1-\lambda-\rho_1\rho_2)(1-\lambda)^{2}} Var(D_1)
\end{aligned}
$$

Thus

$$
\begin{aligned}
Cov(D_{t-1,2}, D_{t-p-1,2}) &= \rho_2^p Var(D_2) - \lambda Cov(D_{t-1,1}, D_{t-p-1,2}) \sum_{i=0}^{p} \rho_2^i \left(\frac{1-\lambda}{\rho_1}\right)^i \\
&= \rho_2^p Var(D_2) - \lambda \sum_{i=0}^{p} \rho_2^i Cov(D_{t-1-i,1}, D_{t-p-1,2}) \\
&= \rho_2^p Var(D_2) + \frac{\lambda^2(1-\lambda)Var(D_1)}{(1-\lambda-\rho_1\rho_2)} \sum_{i=0}^{p-1} \left(\frac{\rho_2^i \rho_1^{p-i+1}}{(1-\lambda)^{p-i+1}}\right)
\end{aligned}
$$

Substitute the above equation to Equation 3.12, we have

$$
\begin{aligned}
Var(q_{t,2}) = &\left(1 + 2\frac{L}{p} + 2\frac{L^2}{p^2}\right) Var(D_2) - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) Cov(D_{t-1,2}, D_{t-p-1,2}) \\
&+ z_2^2 Var(\hat{\sigma}_{et,2}^L - \hat{\sigma}_{et-1,2}^L) \\
= &\left(1 + 2\frac{L}{p} + 2\frac{L^2}{p^2}\right) Var(D_2) - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) \rho_2^p Var(D_2) \\
&- \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) \frac{\lambda^2(1-\lambda)Var(D_1)}{(1-\lambda-\rho_1\rho_2)} \sum_{i=0}^{p-1} \left(\frac{\rho_2^i \rho_1^{p-i+1}}{(1-\lambda)^{p-i+1}}\right) \\
&+ z_2^2 Var(\hat{\sigma}_{et,2}^L - \hat{\sigma}_{et-1,2}^L)
\end{aligned}
\tag{3.13}
$$

Given the above derivation, the bullwhip effects for product 1 and product 2 are:

$$
\frac{Var(q_1)}{Var(D_1)} \geq 1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)\left(1 - \frac{\rho_1^p}{(1-\lambda)^p}\right) \tag{3.14}
$$

$$
\begin{aligned}
\frac{Var(q_2)}{Var(D_2)} \geq\ & 1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)(1 - \rho_2^p) \\
& - \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right) \frac{\lambda^2(1-\lambda)}{(1-\lambda-\rho_1\rho_2)} \frac{Var(D_1)}{Var(D_2)} \sum_{i=0}^{p-1} \left(\frac{\rho_2^i \rho_1^{p-i+1}}{(1-\lambda)^{p-i+1}}\right)
\end{aligned}
\tag{3.15}
$$

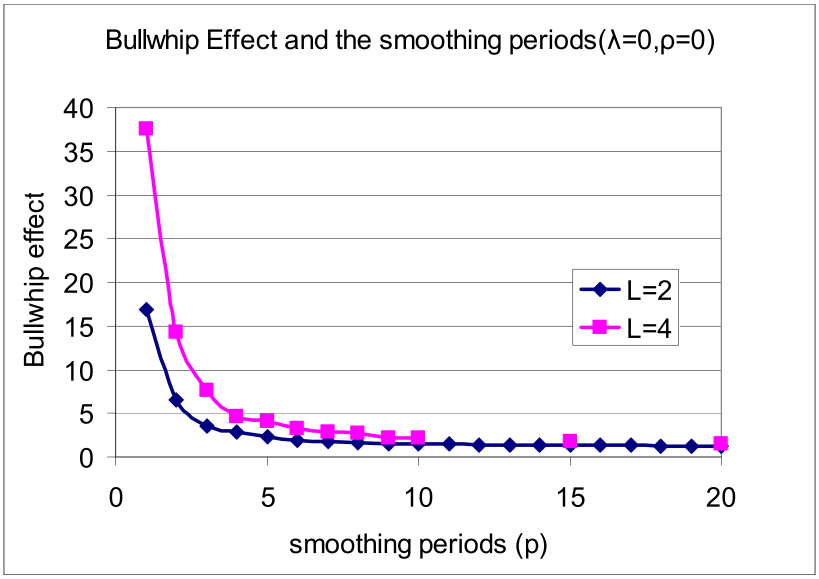


Figure 3.5: Bullwhip effect and the smoothing periods (p)

the bound is tight when $z_1 = 0, z_2 = 0$

3.3 Summary

Several observations can be made from Equation 3.14 and Equation 3.15. First, we notice that the increase in variability from the retailer to the manufacturer is a function of four parameters:(1) p, the number of observations used in the moving average, (2)L, the lead time parameter, (3) ρ_1, ρ_2 the correlation parameter, and (4)λ, the substitution parameter.

The lower limit of the bullwhip effect for both products are a decreasing function of p, the number observations used to estimate the mean and the variance of demand. Figure 3.5 shows the relationship between the bullwhip effect and the smoothing periods (p). When p is large, the lower limit of the bullwhip effect of both product is decreased. However, when p is small, the lower limit of the bullwhip effect can be significantly increased. In other words, the smoother

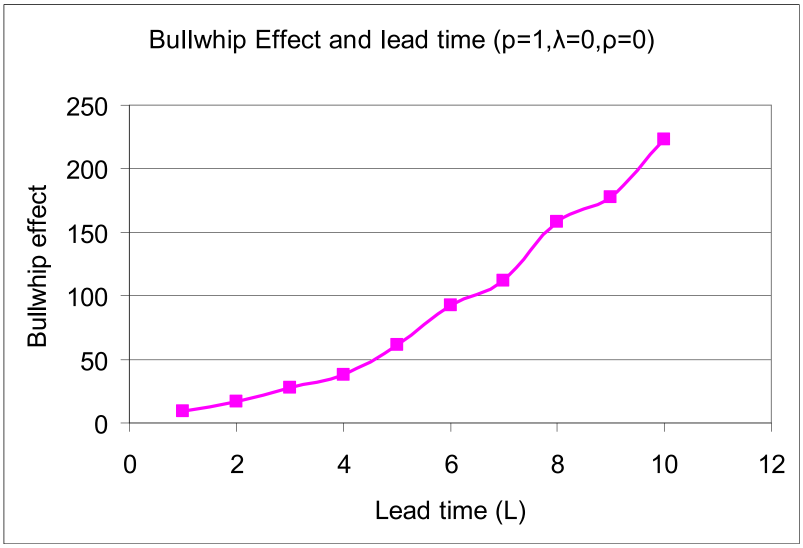


Figure 3.6: Bullwhip effect and the lead time (L)

the demand forecasts, the smaller the increase in the lower limit of the bullwhip effect.

The lower limits of the bullwhip effects of both products are increasing functions of L, the lead time parameter. Figure 3.6 shows the relationship between the bullwhip effect and the lead time L. As can be seen from both Equation 3.14 and Equation 3.15, the larger L, the higher the lower limits are. If the lead times is doubled, to maintain the same order of the lower limit of bullwhip effect, twice demand data must be supplied. That is, the retailer must use more demand data in order to reduce the bullwhip effect if the lead time is longer.

The correlation parameter $\rho_1.\rho_2$ can also play very important role in deciding the lower limit of the variability. If $\rho_1 = 0$ Then Equation 3.14 and Equation 3.15 become

$$\frac{Var(q_1)}{Var(D_1)} \geq 1 + \frac{2L}{p} + \frac{2L^2}{p^2} \tag{3.16}$$

$$\frac{Var(q_2)}{Var(D_2)} \geq 1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)(1 - \rho_2^p) \tag{3.17}$$

If $\rho_2 = 0$,

$$\frac{Var(q_1)}{Var(D_1)} \geq 1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)\left(1 - \frac{\rho_1^p}{(1-\lambda)^p}\right) \tag{3.18}$$

$$\frac{Var(q_2)}{Var(D_2)} \geq 1 + \frac{2L}{p} + \frac{2L^2}{p^2} - \lambda^2\left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)\frac{Var(D_1)}{Var(D_2)}\frac{\rho_1^{p+1}}{(1-\lambda)^{p+1}} \tag{3.19}$$

If $\rho_1 = 0$ and $\rho_2 = 0$

$$\frac{Var(q_1)}{Var(D_1)} \geq 1 + \frac{2L}{p} + \frac{2L^2}{p^2} \tag{3.20}$$

$$\frac{Var(q_2)}{Var(D_2)} \geq 1 + \frac{2L}{p} + \frac{2L^2}{p^2} \tag{3.21}$$

If $\rho_1 \geq 0$, demands are positively correlated, the larger ρ_1, the smaller the lower limit of variability for both products. If $\rho_1 < 0$, demands are negatively correlated, then some strange behavior can be observed. That is, the lower bound of product 1 variability is larger when p is odd values than even value p

If $\rho_2 \geq 0$, demands for product 2 is positively correlated, the variability of product 1 will not be affected. The variability of product 2 will be decreased if ρ_2 is increased. If $rho_2 < 0$, the behavior will be hard to predict since the p value can be even or odd.

The substitution parameter λ is also very important in deciding the lower limit of the variability for both products. If $\lambda = 0$, the lower bound of two products' variability become the same.

$$\frac{Var(q_1)}{Var(D_1)} \geq 1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)(1 - \rho_1^p)$$

$$\frac{Var(q_2)}{Var(D_2)} \geq 1 + \left(\frac{2L}{p} + \frac{2L^2}{p^2}\right)(1 - \rho_2^p)$$

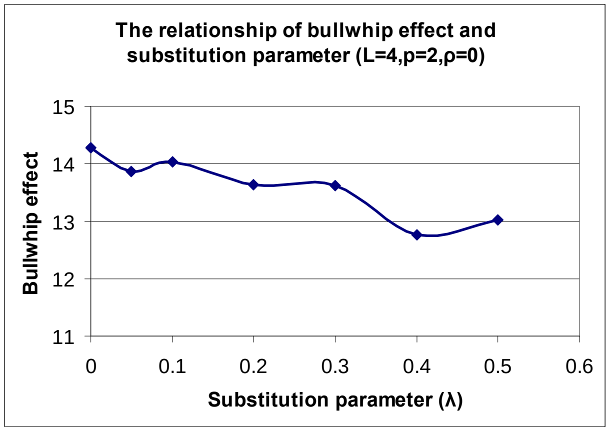


Figure 3.7: The relationship of bullwhip effect and substitution parameter ($L = 4, p = 2, \rho = 0$)

If $\lambda \geq 0$, the larger λ, the smaller the increase in variability for product 1. The increase in variability for product 2 will be decreased with larger λ is $\lambda < 1 - \rho_1\rho_2$. If $\lambda > 1 - \rho_1\rho_2$, then the larger λ, the larger the increase in variability. Notice that, by allowing a fixed percentage of product 1 to substitution product 2, the bullwhip effect for both products are lowered. Figure 3.7 shows the numerical experiments of the bullwhip effect and the substitution parameter λ.

Next we applied the model result to a multistage supply chain network. First we investigate the case that the demand information is shared among all stages. Figure 3.8(a) shows an example of multistage supply chain network. We assume that the network are practicing the same order-up-to inventory policy, the same demand forecast technology, the same number of periods to smoothing demand forecast. The demand substitution is done at stage 1. L^k is the lead time to ship product 1 from stage $k + 1$ to stage k. Since the demand information is shared among all stages, the centralized demand information can be think as increasing of the lead time. Thus the following result can be obtained:

$$\frac{Var(q_1^k)}{Var(D_1)} \geq 1 + 2\left(\frac{\sum_{i=1}^{k} L_1^k}{p} + \frac{\left(\sum_{i=1}^{k} L_1^k\right)^2}{p^2}\right)\left(1 - \frac{\rho_1^p}{(1-\lambda)^p}\right) \tag{3.22}$$

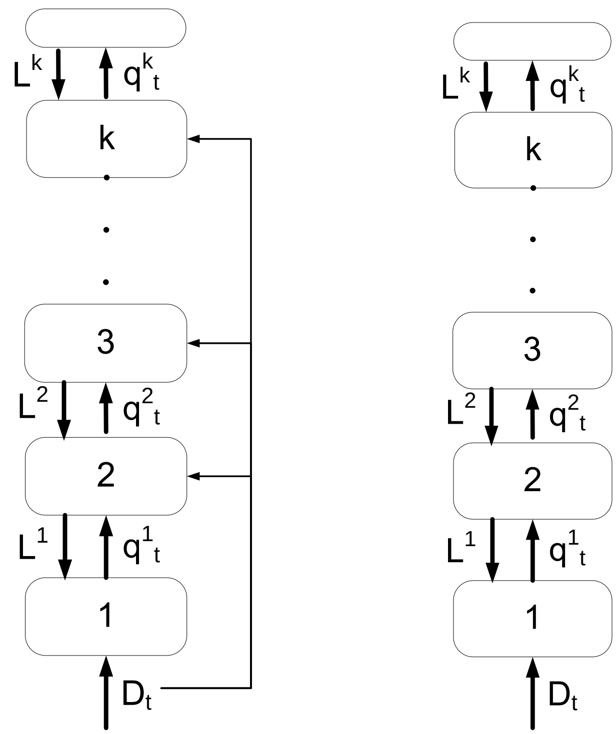
$$\frac{Var(q_2^k)}{Var(D_2)} \geq 1 + 2\left(\frac{\sum_{i=1}^{k} L_2^k}{p} + \frac{\left(\sum_{i=2}^{k} L_2^k\right)^2}{p^2}\right)$$
$$\left(1 - \rho_2^p - \frac{\lambda(1-\lambda)}{1-\lambda-\rho_1\rho_2}\frac{Var(D_1)}{Var(D_2)}\sum_{i=0} p\frac{\rho_2^i\rho_1^{p-i+1}}{(1-\lambda)^{p-i+1}}\right) \tag{3.23}$$

For the case when the demand information is not shared among all stages, the bullwhip is amplified significantly. Figure 3.8(b) shows the supply chain network without sharing the demand information. The demand information is only available to its previous stage. Besides the above assumptions of the centralized case, we assume the the safety coefficient $z_{1,2} = 0$, the substitution parameter $\lambda = 0$, the correlation coefficient $\rho_{1,2} = 0$. Thus the following result can be obtained:

$$\frac{Var(q_1^k)}{Var(D_1)} \geq \prod_i = 1^k \left(1 + \frac{L^i}{p} + \frac{(L^i)^2}{p^2}\right) \tag{3.24}$$

From Equations 3.22, 3.23 and 3.24, centralized demand information can significantly reduce the bullwhip effect.

In this chapter, we also propose a new qualitative measurement, Inventory Entropy, of the bullwhip effect based on the concept of digital signal processing technology. We have compared the traditional bullwhip effect measured by variance and the Inventory entropy. A simulation experiment is conducted. *Order-up-to-S* inventory policy is utilized in the experiment and simple moving average forecast technique is used to forecast the future demand. The periodograms are drawn for both demand process and order quantity pattern. Periodogram shows



(a) Sharing Demand Information

(b) Not Sharing Demand Information

Figure 3.8: A multistage supply chain network with/without shared demand information

not only the amplitude but also the frequency information of the time series data. It shows that our new measurement can accurately describe the bullwhip effect.

Chapter 4

Metamodel Analysis on Multi-Echelon Supply Chain Simulation Experiments

A simulation model is constructed to mimic the reality as closely as possible. With development of the computer technology, large size simulation models for complex systems become possible. However, for very complex systems the simulation models themselves can be large and difficult to understand. Constraints, such as the cost and complexity of the model development, can prevent the modeler from building multiple prototypes of the real system. In such case a simpler model named "*metamodel*" can often be built as a "model of model". The term "metamodel" was introduced by Kleijnen (1986), which has been widely used in the simulation community to study the behavior of complex systems.

Supply chain modeling and simulation is a hot topic in supply chain management research area. Due to its complexity and underlying uncertainty, supply chain is hard if not impossible to be studied in pure mathematical approach.

Only handful mathematical results are available and those results are based on heavily simplified assumptions. It seems like the simulation is the only way to study the system behavior in some cases.

In this chapter, a multi-echelon distribution system with time-based service level requirements is considered. The simulation model is built. We then apply the principles and techniques of metamodel and Design and Analysis of Simulation Experiment (DASE) to reveal the fundamental characteristics. We first present the analysis results of first order single factor polynomial metamodel. Then the multi-order metamodel is presented and discussed. The remainder of the chapter is organized as follows. In section 4.1, the detailed description of the supply chain is given. In section 4.2, the techniques of metamodel analysis is presented. In section 4.3, the analysis results are presented. The conclusion and discussion can be found in section 4.4.

4.1 Multi-Echelon Supply Chain Model

4.1.1 Periodic Review (s, S) Inventory Model

We consider a simple periodic review (s, S) inventory model. If the stock level at the beginning of a period is below a certain value s, then a certain quantity is ordered such that the inventory level reaches S. If the initial stock level is above s, then no replenishment order is placed. Figure 4.1 shows the change of the inventory level and inventory position under a periodic review (s, S) policy with continuous demand where L is the lead time and T is the length of the review period.

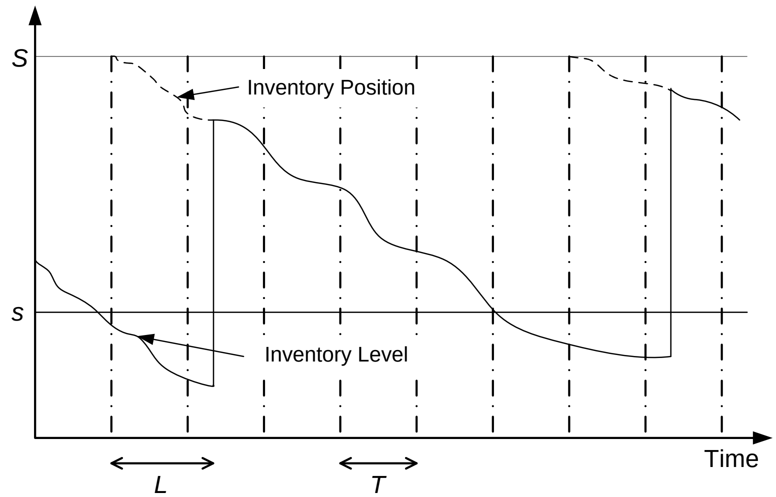


Figure 4.1: A periodic-review (s, S) policy with continuous demand

It has been shown that for a single echelon system the optimal inventory policy is of the (s, S) type (Tijms, 1972). In this chapter, the (s, S) policy is used in all locations of the supply chain.

4.1.2 Model Assumption

The following assumptions are considered:

1. The distribution network has an arborescent structure. That is, each location is replenished from exactly one parent node of higher echelon. There is only one location at the root of the distribution network. Figure 4.2 shows the topology of the distribution network.

2. Demands only occur at the leaf stations in the lowest echelon. All demand locations are assumed to be in the same echelon. To achieve this property, a dummy station might be needed.

3. The demand processes at all locations are independent Poisson processes and the coefficients are known and stationary.

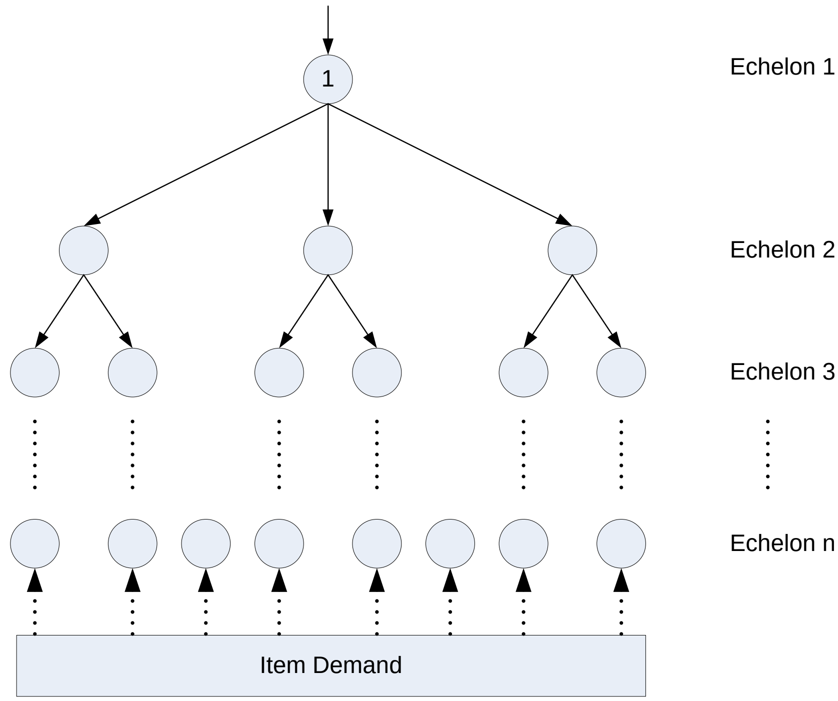


Figure 4.2: The arborescent structure of the distribution network

4. The demand is for one unit of an item at one time.

5. The unsatisfied order will be backordered.

6. The demand location will serve the demand on a first-come-first-serve basis across the distribution network.

4.1.3 The Decision Paradigms

For inventory system research, two principal decision paradigms are used to determine the optimal values of parameters. The first is solely cost driven. A cost function is typically formulated during the modeling process and the optimal value is obtained by using optimal search technology. The cost function assign penalty cost on backlogged demand besides the traditional setup and holding costs. The only objective is to minimize the expected total cost. Many researches have been done using this paradigm, such as Clark and Scarf Clark, Andrew J. and Scarf, Herbert (1960); Hadley and Whitin (1963); Axsäter (2006); Zipkin (2000).

The second paradigm is the service level driven approach. In this paradigm, the constraints of meeting the customer demands within a certain probability are considered besides the cost function. Customer demands are either satisfied immediately or in a short time period. Silver et al. (1998) discusses a list of the service level measurements. Two commonly used definitions are the probability of no shortage and the probability of the fill rate. Although the cost driven paradigm is straightforward, the assessment of the penalty cost, particulary the cost of losing customer goodwill is very difficult. Thus service level driven approaches are very popular in practice. Under this approach, the objective is to obtain an optimal value pair of (s, S) that minimizes a cost function which only includes setup and holding costs, subjected to the constraint that the solution satisfies a predetermined customer service level. In this chapter, the service level is measured by the fill rate, which is defined as the probability that an arriving demand for an item will be satisfied within a specified period of time.

4.1.4 Model Configuration

Figure 4.3 shows the supply chain considered in this chapter. There are three echelons. Echelon 1 has only one station “1”. Station “1” is replenished from echelon 0 (not shown in this figure). The lead time for the replenishment is 5 days. Echelon 2 has two stations “2” and “6”. They are replenished from station “1”. The lead time between echelon 1 and 2 is 2 days. Echelon 3 has 6 stations. Stations “3”, “4” and “5” are replenished from station “2”; stations “7”, “8” and “9” are replenished from station “6”. The lead time between echelon

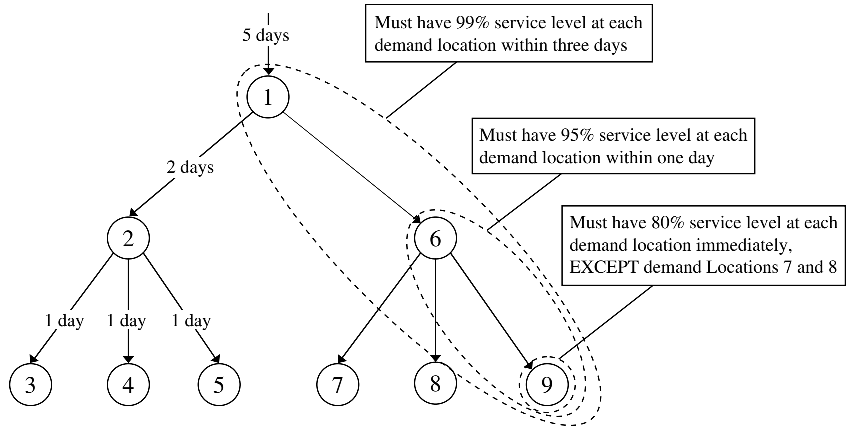


Figure 4.3: A multi-echelon supply chain network

2 and echelon 3 is 1 day. The service level requirements are different among different demand stations. Demand stations "3", "4", "5", and "9" require 80% immediate service level, 95% service level within one day and 99% service level within 3 days. Demand stations "7", "8" have the same one-day and three-day requirements, but no requirement on immediate service. When the demand can not be satisfied immediately, we impose a higher shortage cost. The question of interest is how to reduce the average total cost.

$$C_{\text{Avg Total}} = C_{\text{Avg ordering}} + C_{\text{Avg shortage}} \tag{4.1}$$

4.2 Metamodel Analysis

The mathematical representation of a simulation's input-output function can be written as

$$\bar{Y} = f(\bar{X}) \tag{4.2}$$

where $\bar{Y}, \bar{X}$ are vectors and will usually contain random components. In our simulation model, the vector $\bar{X}$ will include the demand rates, s and S for each station and the initial inventory for each station. The lead time and the transportation time might also be included. $\bar{Y}$ could be the average total cost, average holding time, average shortage cost and so on. The response vector $\bar{Y}$ might have more than one value, but we typically focus our attention on only one component of $\bar{Y}$. If so, Equation (4.2) can be written as

$$y = f(\bar{X}) + \epsilon \tag{4.3}$$

Then the task of metamodel is to find a good approximation of function f and the model of ϵ. The major issues in metamodel include i) the choice of a functional form for f, ii) the design of experiments, i.e.. the selection of a set of x points at which to observe y (run the full model), the assignment of random number streams, the length of runs, et al., and iii) the assessment of the adequacy of the fitted metamodel (confidence intervals, hypothesis tests, lack of fit and other model diagnostics). The functional form will generally be described as a linear combination of basis functions from a parametric family. So there are choices for families (e.g., polynomials, sine functions, piecewise polynomials, wavelets, etc.) and choices for the way to pick the representation horn within a family (e.g. least squares, maximum likelihood, cross validation, etc.). The issues of experiment design and metamodel assessment are related since the selection of an experiment design will be determined in part by its effect on assessment issues (Barton, 1992).

4.2.1 Mathematical Model

As stated previously, a metamodel is used to find the input-output relationship. Given this purpose, the inputs are carefully designed based on some principles. For details of the design of the input value please refer to Kleijnen (2008). The experiments are carried based on the designed input values. The output data are collected as follows:

$$Y_{ijk} : i = 1, ..., N; j = 1, ..., R_i; k = 1, ..., K_{ij}$$

where N is the number of experiment conditions. R_i is the number of replications of experiment condition i. K_{ij} is the total number of observation for experiment condition i, replication j.

The mean response for each replication is

$$\bar{Y}_{ij\cdot} = \sum_{k=1}^{K_{ij}} \frac{Y_{ijk}}{K_{ij}}, i = 1, ..., N, j = 1, ..., R_i$$

The central limit theory indicates that when the sample size is large, the output approximates normal distribution.

$$\bar{Y}_{ij\cdot} \sim N(\mu_i, \sigma_i^2) \tag{4.4}$$

If the target response is the sum or average of certain observations, the target can be modeled using the normal distribution based on the *Central Limit Theorem*. The Normal distribution can be adequately described by its mean and variance parameters. Now the parameters μ_i, σ_i^2 needs to be estimated based on the collected data.

Many statistical methods are available to estimate the parameters, but most of them are based on the assumption that the observations are collected and independent and identically distributed (*i.i.d.)*. One common technique is to test the following hypothesis.

$$
\begin{aligned}
H_o &: \beta_1 = 0 \\
H_1 &: \beta_1 \neq 0
\end{aligned}
\tag{4.5}
$$

where β_1 is the coefficient of the first degree polynomial regression on the output:

$$W_i = \beta_0 + \beta_1 i + \epsilon$$

if β_1 is sufficiently different from zero, then the *i.i.d.* assumption is not held for the observations. The positive β_1 corresponds to the learning curve and the negative value corresponds to the fatigue curve (Leemis, 2004).

For the observation collected from *i.i.d.*, the parameters can be estimated using Maximum Likelihood Estimators (MLE), Least Squares and the method of moments. MLEs are the most used in this chapter due to special properties. MLEs for the normal distribution are

$$\hat{\mu}_i = \bar{Y}_{i\cdot\cdot} = \frac{1}{R_i}\sum_{j=1}^{R_i} \bar{Y}_{ij\cdot} \ , \ \hat{\sigma}_i = \left(\frac{1}{R_i}\sum_{j=1}^{R_i}(\bar{Y}_{ij\cdot} - \bar{\bar{Y}}_{i\cdot\cdot})^2\right)^{1/2}$$

After a parameter is estimated, the statistical *goodness of fit* tests like Chi-square and Kolmogorov-Smirnov tests are used to investigate the quality of the estimation. The effectiveness of metamodel can also be justified by using test data. The results are often compared with outputs from corresponding simulation model. In this chapter the metamodel is validated by using the test data set.

If the parameter is estimated by MLEs. The following confidence interval can be used to test the effectiveness of the metamodel (Law and Kelton, 2000): $I_i(\mu) = \hat{\mu}_i \pm \sigma_i(\mu)$ and $I_i(\sigma) = \hat{\sigma}_i \pm \delta_i(\sigma)$ where

$$\delta_i(\eta) = \left(-E\left[\frac{\partial^2 \ln L_i(\hat{\mu}_i, \hat{\sigma}_i)}{\partial \eta^2} \right] \right)^{-1/2} \tag{4.6}$$

and

$$\ln L_i(\mu_i, \sigma_i^2) = -R_i \ln \sqrt{2\pi} - \frac{R_i}{2} \ln \sigma_i^2 - \frac{1}{2\sigma_i^2} \sum_{j=1}^{R_i} (\bar{Y}_{ij\cdot} - \mu_i)$$

If the metamodel's output with the estimated parameters is within the confidence interval, the metamodel is not rejected. For example, if 95% of the experiment conditions are within the confidence interval then the metamodel is said not to be rejected.

$$H_0 : g_1(X_i; \hat{\theta}_\mu) \in I_i(\mu) \cap g_2(X_p; \hat{\theta}_\sigma) \in I_i(\sigma) \tag{4.7}$$

Where g_1 and g_2 are the transformation functions of the mean and variance.

4.3 Simulation and Results

4.3.1 Single Factor Metamodel

We build a simulation model to gain insights of the behavior of the supply chain so that the average total cost can be reduced. Thus the output here will be the average total cost. Many decision variables can be identified from Figure 4.3.

We are interested in the initial inventory for each station, the selection of small s and big S. We assume that the demand pattern at each location in the lowest echelon is the same. That is, the demand inter-arrival times are the same across all demand locations. The demand inter-arrival time ranged from 0.1 to 2 corresponding to the busy and idle situation respectively. We divide the range into 10 smaller ranges at the following point 0.1, 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9. Given the demand inter-arrival time, the initial inventory and the reorder level s and order up to level S need to be decided. Only after these variables are determined, the average total cost can be obtained by simulation. In other words, the final average total cost is affected by the demand inter-arrival time, the initial inventory, and the selection s and S. The model was developed using Arena 10.0 and the data were analyzed using MATLAB 2000. Figure 4.4 shows the screen shot of a running model. For each experiment condition, we replicated for 100 times independently. For each replication, 2000 observations were collected. The simulation length was set to be 300 time units (day) to reach the steady output of the average total cost.

Figure 4.5 shows the relationship between the average total cost and the simulation time. At the beginning of the simulation, the fluctuation of the average total cost is large. Then it becomes much stabler. To better represent the simulation output, the data points at the beginning of each replication will be removed due to the high fluctuation. The Welch's moving average (Law and Kelton, 2000) algorithm will be used to identify the cutoff point.

The initial inventory of each location is determined by using Equation (4.8).

$$I = d(T + L) + z\sigma_d\sqrt{T + L} \tag{4.8}$$

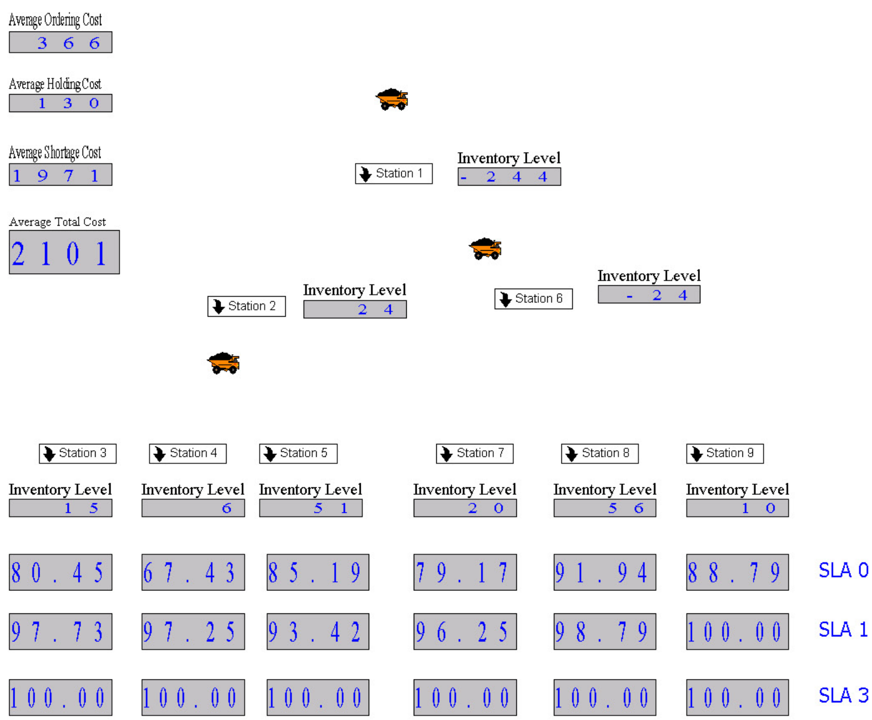


Figure 4.4: The screenshot of the running model

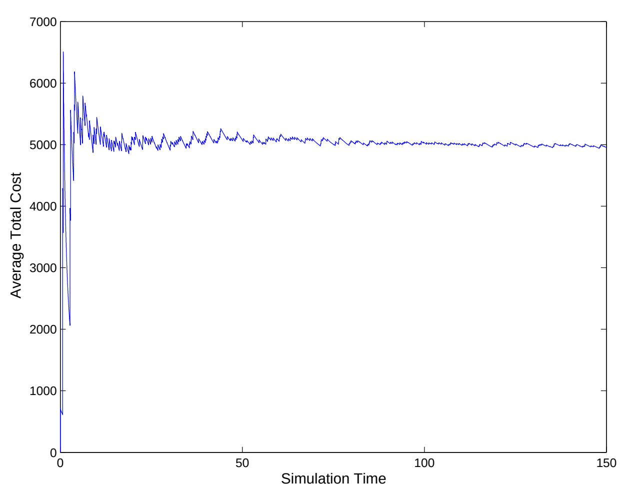


Figure 4.5: The average total cost versus the simulation time (demand interarrival time is 0.5)

where I is the calculated initial inventory and d is the average daily demand. If the location is not in the lowest echelon then the demand will be the sum of all its children's demand; T is the length of the review interval; L is the lead time; σ_d is the standard deviation of the demand per time period ; and z is the safety factor (Sahin and Robinson, 2006). $\Phi(z) = b/(b+h)$ where b is the shortage cost and h is the holding cost. In our model, we set $b = 10, h = 1$. Given the demand rate is 0.5, the calculation of the initial inventory of the location 2 is as follows:

$$\begin{aligned}
d &= 1/0.5 + 1/0.5 + 1/0.5 = 6 \\
T &= 1 \;\; L = 2 \\
\sigma_d &= 6 \\
z &= 1.335 \;\; \Phi(z) = \frac{10}{10+11} \\
Q &= 6 * (1+2) + 1.335 * 6 * \sqrt{1+2} = 31.87
\end{aligned}$$

We round the calculated value to the nearest integers towards infinity. The initial inventory for location "2" is 32. Since the demand rates are the same across all locations in the lowest echelon, we assume the initial inventory is the same on each echelon. Note that Equation (4.8) is actually the *order-up-to* level of the periodic review system with on-hand inventory is zero. The main purpose of the initial inventory is to meet the service level agreement especially the immediately service level agreement, denoted by SLA0, at the beginning of the simulation. Location "7" and "8" have no requirement on immediate service, thus this calculation will tend to give more initial inventory than the real need. But we will simply assume they are the same since the difference is not big.

Table 4.1: The initial inventory level and the optimized s and S. The subscript means the echelon level.

Demand interarrival time	Inv_1 S_1	Inv_2 S_2	Inv_3 S_3	s_1	s_2	s_3
0.1	557	160	39	141	71	24
0.3	186	54	13	47	24	8
0.5	112	32	8	29	15	5
0.7	80	23	6	21	11	4
0.9	62	18	5	16	8	3
1.1	51	15	4	13	7	3
1.3	43	13	3	11	6	2
1.5	38	13	3	10	5	2
1.7	33	10	3	9	5	2
1.9	30	9	3	8	4	2

We calculate S in the same way as initial inventory based on Equation (4.8). s is calculated using Equation (4.9) and rounded to the nearest integers towards infinity.

$$s = dT + z\sigma\sqrt{T} \tag{4.9}$$

We also assume that the s and S are the same on each echelon. That is, locations $3, 4, 5, 7, 8, 9$ will have the same s and S, while location 2 and 6 will have the same s and S. Table 4.1 shows the the results of the initial inventory and the corresponding s and S. The search for optimal s and S could also be accomplished by using the application named *OptQuest* from Arena.

We choose $N = 10, R_i = 100, K_{ij} = 2000$ and collect data to build our metamodel. Table 4.2 shows the Maximum Likelihood Estimations, the $\sigma_i(\eta)$. As it can be seen from the table, the value of β_1 is very close to 0. That is, the assumption (4.5) is valid for the collected data.

Table 4.2: Experiment condition, the MLE estimators, $\delta_i(\eta)$

i	X_i	β_1	$\hat{\mu}_i$	$\delta_i(\mu)$	$\hat{\sigma}_i$	$\delta_p(\sigma)$
1	0.1	0.212	1710.750	7.286	36.538	5.208
2	0.3	-0.002	676.125	2.895	14.517	2.069
3	0.5	-0.065	463.519	2.214	11.101	1.582
4	0.7	-0.069	355.158	2.050	10.278	1.465
5	0.9	-0.034	324.532	1.612	8.084	1.152
6	1.1	-0.043	259.489	1.452	7.281	1.038
7	1.3	-0.047	249.443	1.436	7.200	1.026
8	1.5	-0.050	227.221	1.463	7.338	1.046
9	1.7	-0.030	205.267	1.344	6.742	0.961
10	1.9	-0.026	190.939	1.294	6.490	0.925

Table 4.3: Number of rejected points versus the degree of the polynomial regression

degree	1	2	3	4	5	6	7	8	9	10
mean	10	10	10	10	10	8	5	6	0	0
std	10	7	6	4	0	0	0	0	0	0

The polynomial linear regression algorithm is used to find the relationship between the experiment condition (mean demand inter arrival time) and the simulation output (average total cost). The degree of the polynomial regression is important. Higher degree will offer better fit but it tends to overfit. We will test the degree from 1 to 10. The number of points will be recorded if the output of regression is out of the Maximum Likelihood Estimation Confidence Interval. The results are shown in Table 4.3. From the table, we know that the best polynomial degree of the mean average total cost is 9 and the best degree for the standard deviation is 5. Figure 4.6 and 4.7 show the estimated value of the mean and the standard deviation of the average total cost.

The result of the regression is shown in Table 4.4. Figures 4.8 and 4.9 show the value calculated from the metamodel and the value obtained through the simulation. As it can be seen from the figures, the prediction is fairly accurate.

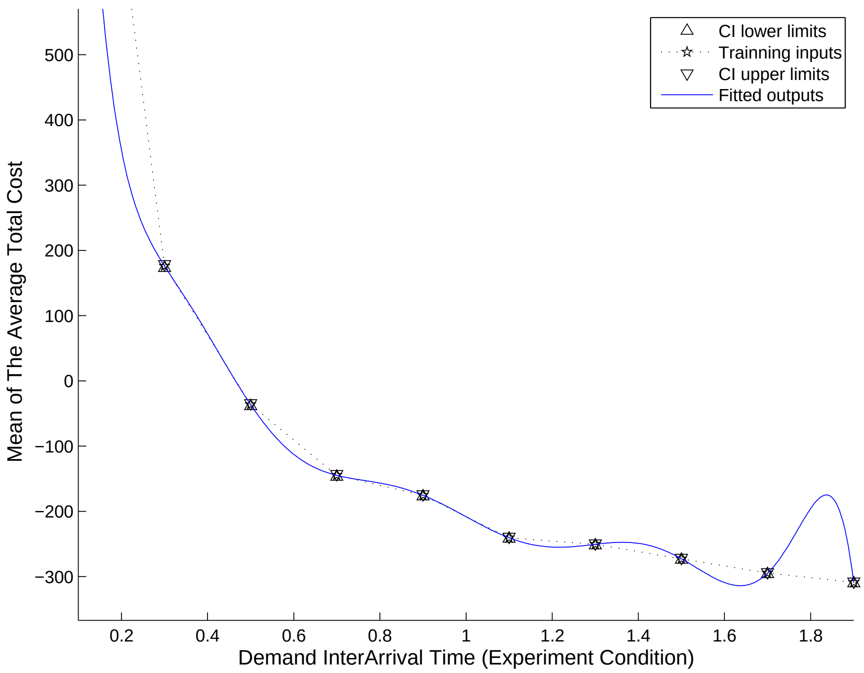


Figure 4.6: Estimated mean average total cost with the respective confidence intervals. The solid line is the polynomial fitting line with degree 9.

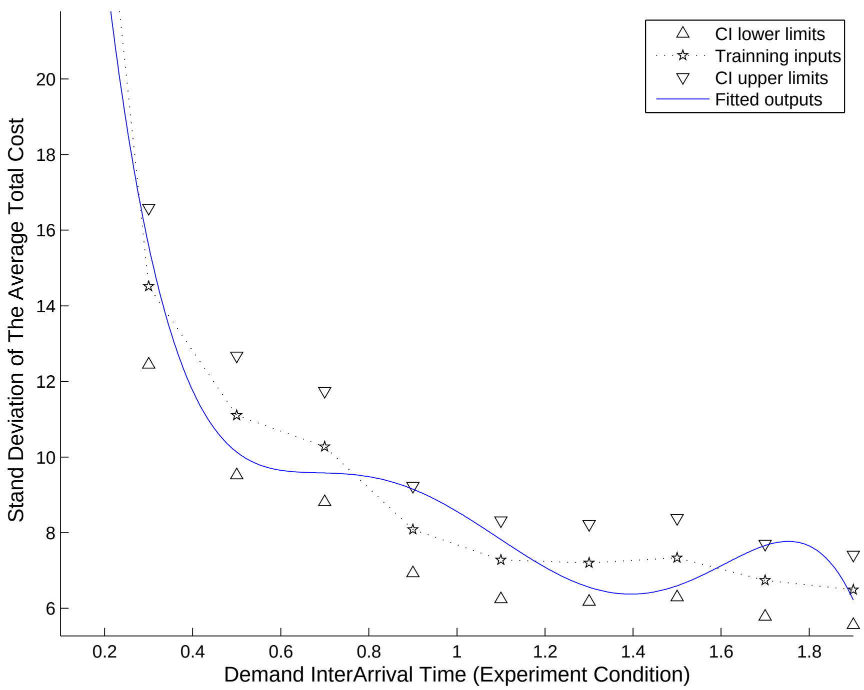


Figure 4.7: Estimated standard deviation of the average total cost with the respective confidence intervals. The solid line is the polynomial fitting line with degree 5.

Table 4.4: Polynomial Regression parameters (The Metamodel)

	mean	stddev
X^0	4.765 E03	57.425
X^1	-6.222 E04	-261.499
X^2	3.648 E05	549.541
X^3	-1.184 E06	-548.750
X^4	2.303 E06	257.470
X^5	-2.788 E06	-45.624
X^6	2.116 E06	
X^7	-9.775 E05	
X^8	2.511 E05	
X^9	-2.747 E04	

Our metamodel shown in Table 4.4 is a good approximation of the real simulation model at the demand interarrival time range from 0.1 to 1.9. Figure 4.6 shows that at the end of this range the metamodel can not accurately predict the mean of the average total cost. Large fluctuation can be observed there. This might indicate that a new metamodel is needed to study the behavior for range greater than 1.9.

4.3.2 Multi Factors Metamodel

In the previous section, the single factor metamodel is presented and discussed. In this section, we will present the multi factors metamodel. In the real world, the supply chain performance is influenced by many factors. The inter-arrival rate is just one of the important factors. Many factors can be identified easily from the model. Below are some examples:

- Order Setup Cost
- Inter-Arrival Rate
- Unit Purchase Cost

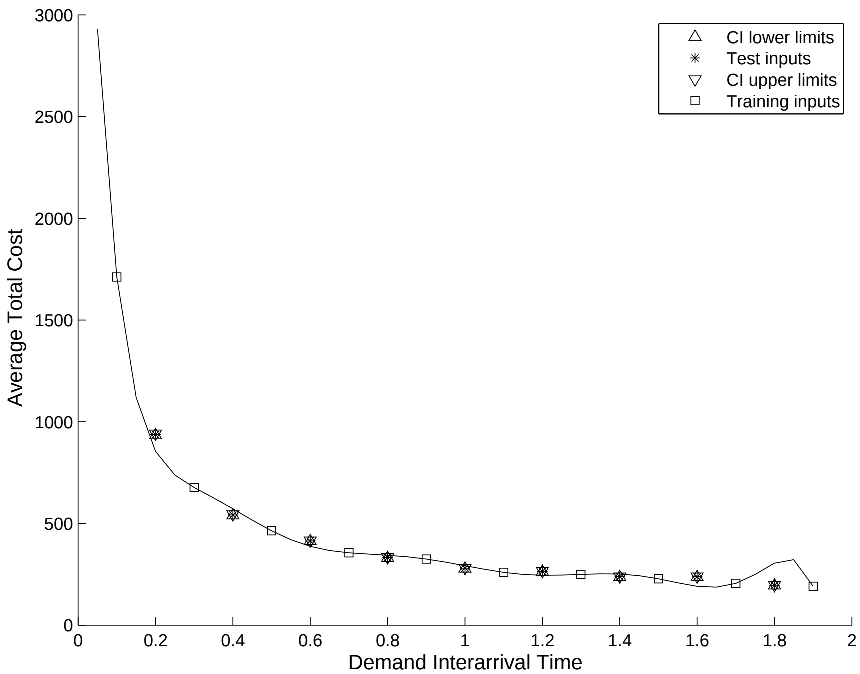


Figure 4.8: The predicted mean value on the test demand interarrival time versus the fitted polynomial line with degree 9.

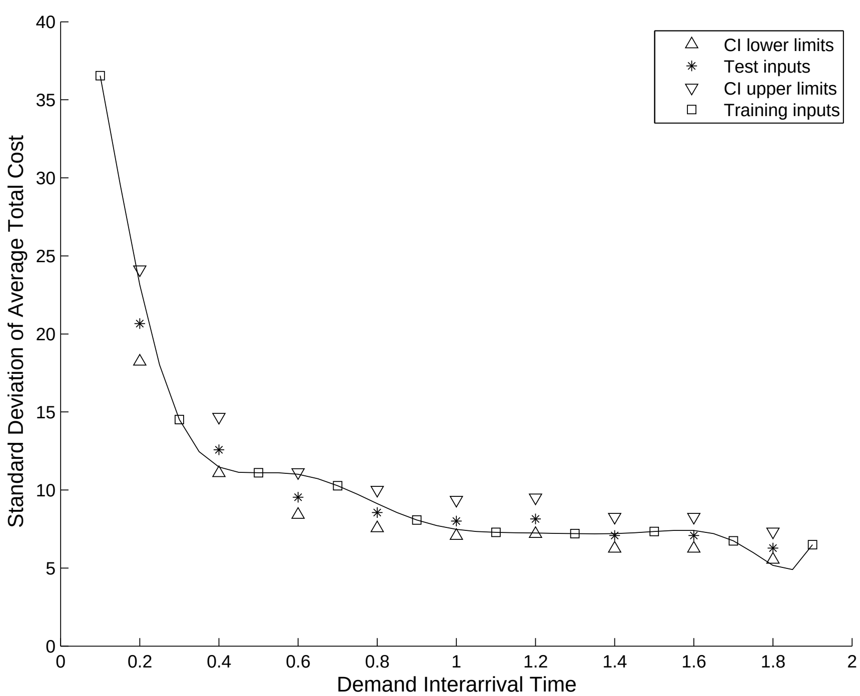


Figure 4.9: The predicted standard deviation value on the test demand interarrival time versus the fitted polynomial line with degree 5.

- Shortage Cost
- Interest Rate
- Initial Inventory for Echelon 3
- Initial Inventory for Echelon 2
- s for the (s, S)Policy
- S for the (s, S)Policy

If the (s, S) policy is different for different station, the number of variables is even more. The large amount of factors bring problem to the simulation study even for computer simulation. Considering 10 factors, each factor has 4 possible values. For each design points, 5 replicates are run. Then the total number of replication is $4^{10} * 5 = 5,242,880$. For large systems, the actual simulation time will also be long. Such high number of replicates will prevent even computer simulation from running complete experiments. Thus the principle of DASE must be applied.

To simplify the research of this supply chain model, the following modifications are made on the simulation configuration:

- Echelon 1 will be out of our consideration. The cost in Echelon 1 will not be calculated.
- All the stations in Echelon 3 will have the same service level agreement. Which means station 7 and 8 will have the same SLAs as the other stations in echelon 3
- The parameters are the same across the same echelon. Thus all station in the same echelon will have the same s and S.

Table 4.5: High Value and Low Value of Each Factor

Name	Low Value	High Value
s	2	29
S	40	90
InitInv2	3	20
InitInv3	10	60
Interest Rate	0.01	0.10
Inter Arrival	0.10	1.00
SetupCost	30	60
ShortageCost	10	20
UnitCost	30	60

After having these assumptions and the assumptions mentioned in the previous section, the following 9 variables are identified: s, S, Unit Cost, Inter-Arrival Rate, Initial Inventory Echelon 2, Initial Inventory Echelon 3, Interest Rate,Shortage Cost, and Setup Cost. The objective is still the same: minimize the average total cost. The constraints are to satisfy the SLAs.

The following analysis is based on 2^9 factorial design. We arbitrarily select two values of each factor. Table 4.5 shows our selected values for each factor. Figure 4.10 shows the relationship between the s,S and the SLAs. Larger values of s and S tend to have high SLA0 value. That is, the possibility of satisfying the customer instantly is high. But high values of s and S mean high holding cost and order cost generally.

By using DASE methodology the following metamodel can be obtained from the simulation model.

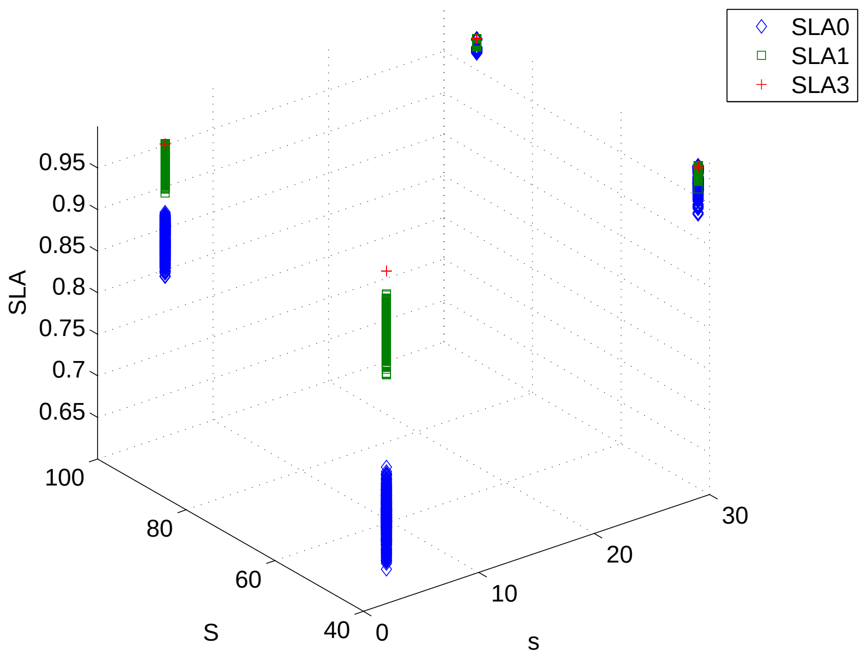


Figure 4.10: The relationship among s, S and the SLAs (Service Level Agreements). SLA0 is the percentage of instantly satisfied customer, SLA1 is the percentage of customers who wait less than one day, including 0 day, SLA3 is the percentage of customers who wait less than 3 day.

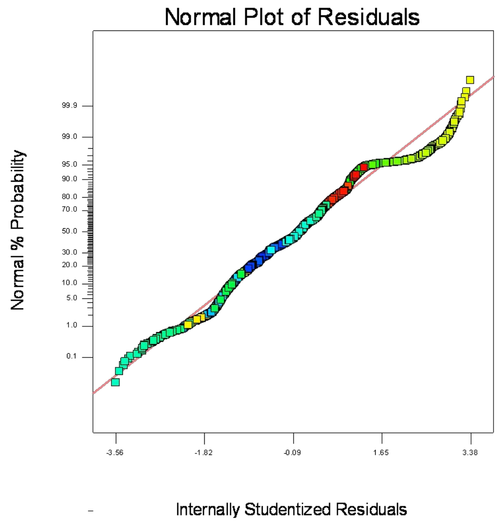


Figure 4.11: The normal plot of residuals

$$
\begin{aligned}
AverageTotalCost &= +16.636 + 13.297 * s \\
&\quad -0.845 * S - 21748.149 * InterestRate \\
&\quad -155.397 * InterArrival + 59.968 * UnitCost \\
&\quad +334.945 * S * InterestRate \\
&\quad +438.932 * InterestRate * UnitCost \\
&\quad -48.619 * InterArrival * UnitCost
\end{aligned}
\tag{4.10}
$$

Equation 4.10 shows the relationship between the Average cost and the s, S, Interest Rate, Inter Arrival Rate and Unit Cost. Figure 4.11 shows the residual of the above metamodel. $R^2 = 0.9497$ and $R^2_{adj} = 0.9498$ combining with Figure 4.11 prove that the model is significant. The metamodel shows that the average cost is mainly affected by s, S, Interest Rate, Unit Cost, Inter Arrival Rate and their second order intersection term. Table 4.6 shows the contribution of each factor.

Table 4.6: Factor Contributions

Factors	%Contribution
InterArrival	32.93
Interest Rate	23.45
Unit Cost	21.93
S	5.72
$S*$Interest Rate	4.21
InterArrival$*$UnitCost	3.19
Interest Rate$*$UnitCost	2.60
other terms	5.01
s	0.95

The Inter Arrival rate is the most important factor among all factor. The increasing inter arrival rate (decreasing average number of customers arrived per day) will lead to decreased average total cost since there is no need to maintain a high level of stock and order. For an efficient supply chain, the ability to adjust with the customer inter arrival rate is important. Fail to do so will lead to increasing inefficiency of inventory. Term $S*InterestRate$ and $UnitCost*InterestRate$ are clearly related with holding cost. A large S or $UnitCost$ will lead to high holding cost. The shortage cost is not shown in the metamodel because of the high SLA. The required instant service level agreement is over 80% which means the possibility of shortage is only 20%. Thus in the experiment data, the shortage cost will not affect the total cost significantly. Practical policy can be inferred from the above metamodel: to have lower average total cost, s needs to be low, S needs to be large to reduce the ordering cost and purchase cost.

4.3.3 Metamodel Analysis on Products Substitution Case

Demand Substitution is hard to be mathematically formulated due to its inherent stochastic inventory changing process. There are many ways to tackle this difficulty. Metalmodel analysis can be used to formulate the relationship between the input variable and output variable. In this section, a demand substitution model will be analyzed using metamodel methodology.

We assume the demand process following the proportion rule. That is the customer purchases are rationed according to a proportion rule when stock-outs occurs. The proportion rule is used in many literatures Hopp and Xu (2008); Netessine and Rudi (2003). We then assume that the customer didn't know the inventory information in the beginning. When he experiences stock-out, the inventory information of all products is provided to customer so that the customer can make decision accordingly. Since the inventory information is known after the first try, the customer will not pick any stock-outs item. Thus only one substitution is considered in this case.

We consider three products in a market with the product attraction numbers "$1, 1, 1, 1$". The first item in the set is the option of leaving. The customer may choose to leave instead of buying any product. To facilitate the model formulation, we consider the customer leaving the same as products. The last three items are the production attraction numbers of product $1, 2, 3$. The initial inventory levels are set as product $2, 250$, product $3, 300$. Product 1's inventory varies from 150 to 250. The number of customer arrivals follows Poisson distribution with mean 1000. Because the same product attraction number of each product, the average demand for each product is 250. Thus, product 3 is likely to have leftovers and product 1 will experiencing stock-out.

Hopp and Xu (2008) proved that the relationship between service rate and the inventory level is a one to one relationship. We are interested in obtaining the relationship between the service rate and inventory position. The inventory of product 2 and product 3 are kept constant. Only the inventory of product 1 is changed from 150 to 250.

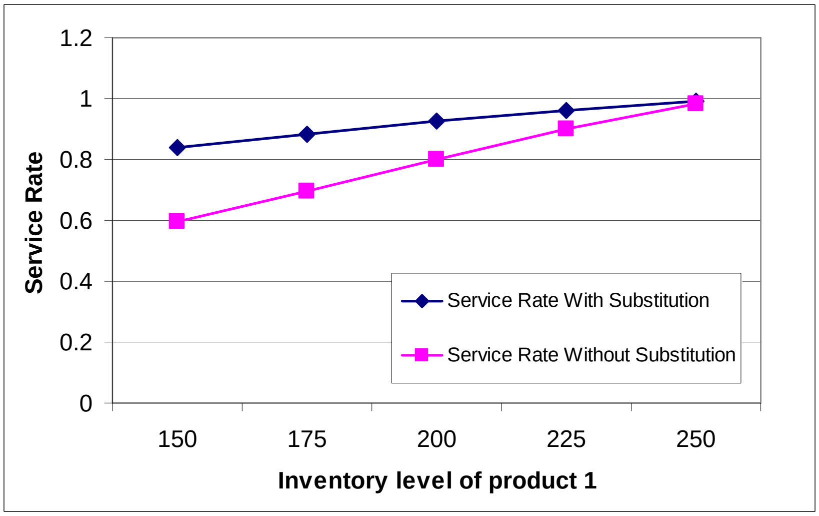


Figure 4.12: The relationship of service rate and inventory of product 1

Figure 4.12 shows the relationship between the service rate and the inventory level of product 1. Table 4.7 shows the result of the simulation for the product substitution case. The result of the Metamodel is given by Equation 4.11.It is clearly shown that the service rate is always higher when the product substitution is permitted. This suggests that ignoring demand substitution may significantly underestimate the effective demand and result in inappropriate inventory and pricing policy, which will eventually hurt the overall profit.

$$\text{Service Rate} = 0.1479 * \text{Inventory}/100 + 0.6281 \tag{4.11}$$

4.4 Summary

In this chapter, we use the methodology of metamodel to investigate a multi-echelon supply chain network with time based service level agreement. Simulation models are built to study the system behavior and DASE techniques are applied to obtain insights and practical findings.

Table 4.7: The service rate of the simulation model with substitution

Inventory	Average	Standard Deviation	#Replication
150	0.8453	0.0312	10
175	0.8897	0.0277	10
200	0.9271	0.0191	10
225	0.9637	0.0208	10
250	0.9932	0.0166	10

Metamodels are abstractions of the simulation model that exposes the system's input-output relationship through simple mathematical functions. We look into both single factor metamodel and multi-factors metamodel for a multi-echelon supply chain simulation model. For the single-factor metamodel, the demand inter-arrival rate is set to be the input variable and the average total cost is set to be the output variable. The goal is to reduce the average total cost as much as possible. To calculate the average total cost, the initial inventory and the s and S need to be calculated first. The initial inventory and S are calculated according to the *order-up-to* level of the periodical review model. The value of s is calculated based on the results of continuous review model. The metamodel is tested by applying to the testing samples. The test results show that the metamodel performs well in the range from 0.1 to 1.9. The fitted closed-form expressions for the mean and standard deviation of average total cost are obtained. These expressions can be used as surrogate models to substitute the actual simulation model in its parent model.

For multi-factor metamodel, nine variables are incorporated in the metamodel, DASE analysis shows that the resulting metamodel is significant. Practical policies are inferred from the metamodel. Although this study focuses on the illustrative multi-echelon supply chain problem, we deem that the metamodel methodology, coupled with DASE techniques, can find wide applications in many other complicated systems.

Chapter 5

Stochastic Network Representation of Demand Substitution Process

Mahajan and van Ryzin (2001b) proposed a fluid model to model the inventory process. The inventory is considered as a fluid and each customer requires a continuous quantity of fluid. The choices are ordered based on the customer's preference vector. The customer drains the inventory of the most preferred fluid first. If this fluid runs out, the customer drains the second preferred inventory and so on. This process ends when either the customers requirement is met or the inventory fluid is exhausted. If each customer requires exactly one unit and the fluid level are integral then the model is a discrete inventory model. Hopp and Xu (2008) utilize this model and use static approximation of the demand substitution. In this research, we will use the similar settings.

We consider a market with L products, indexed by $L = \{1, 2, \ldots, l\}$, and assume that customers have a no-purchase option: 0. That market is open during a finite time interval [0,T], called a season and begins with initial inventory levels y_i of product $i = 1, 2, \ldots, l$. The inventory is assumed to be replenished at the beginning of each season. There is no inventory replenishment during the season. We denote $\chi_i(t)$ as the indicator of the availability of product i at time t; that is $\chi_i(t) = 1$ if $y_i(t) > 0$; other wise $\chi_i(t) = 0$, where $y_i(t)$ is the inventory level of product i at time t and $y_i(0) = y_i$. Hence, if a customer picks product i at time t, and the product is in stock, the inventory level of product i drops by 1, that is $y_i(t+) = y_i(t) - 1$, and χ_i drops to 0 from 1 when product i stocks out. We let $\aleph(t) = \{i \in L | \chi_i(t) = 1\}$ be the set of available products at time t. It is obvious that $\aleph(0) = L$ and $\aleph(s) \subseteq \aleph(t)$, where $s > t$.

The customer arrivals follows an exogenous stochastic process like Poisson process that is independent of the set of available products. At time t a customer selects product i from the set of products that are still available with probability

$$P_{i|\aleph(t)} = \frac{r_i}{r_o + \sum_{j \in \aleph(t)} r_j} \tag{5.1}$$

where $r_i, i = 0, 1, \ldots, l$ are assumed to be independent of the product availability and time. Hence, r_i is the attraction factor for product i which may depend on the price, quality and other attributes of product i. In other words, the customer is not aware of the availability of any product. The customer will pick product according to his own preference which is not affected by the availability of product.

Given the availability process $\{\chi_i(t), t \in [0, T]\}, i \in L$, we can calculate the expected effective demand for product i as

$$E\left[\int_0^T \frac{r_i \chi_i(t)}{r_0 + \sum_{j=1}^{L} r_j \chi_j(t)} dN(t)\right] \tag{5.2}$$

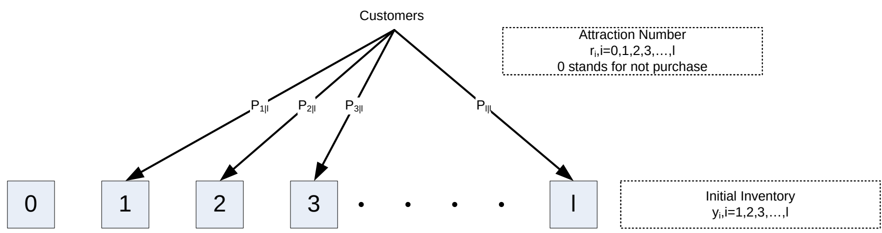
where $N(t)$ is the exogenous customer arrival process. We denote c_i as the unit cost of product i, p_i as the price, and v_i as the salvage value and assume that

$p_i > c_i > v_i$. Then the expected profit for product i is

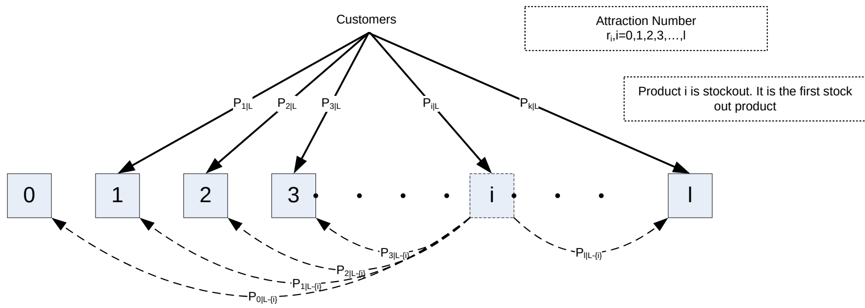
$$
\begin{aligned}
& p_i E\left[\int_0^T \frac{r_i \chi_i(t)}{r_0 + \sum_{j=1}^{L} r_j \chi_j(t)} dN(t)\right] \\
& + v_i \left[y_i - E[\int_0^T \frac{r_i \chi_i(t)}{r_0 + \sum_{j=1}^{L} r_j \chi_j(t)} dN(t)] \right] \\
& - c_i y_i \\
& = (p_i - v_i) E\left[\int_0^T \frac{r_i \chi_i(t)}{r_0 + \sum_{j=1}^{L} r_j \chi_j(t)} dN(t)\right] - (c_i - v_i) y_i
\end{aligned}
\tag{5.3}
$$

Due to the demand substitution effect, the product availability process $\chi_i(t)$ of product i is jointly determined by inventory levels of all products, which results in a major technical difficulty.

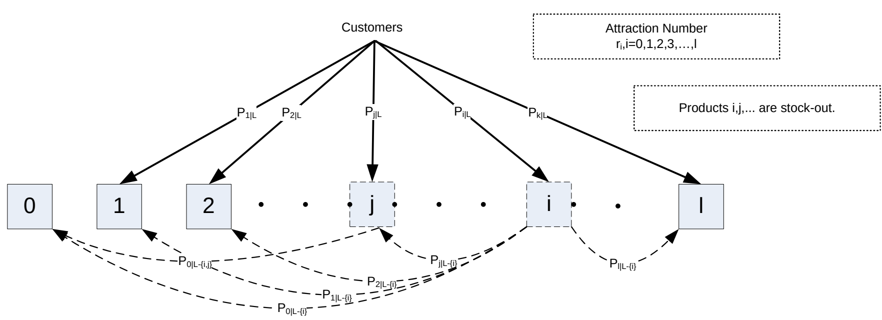
In Figure 5.1(a), the initial state of the substitution model is shown. The initial inventory for each product is $y_i, i = 1, 2, \ldots, l$. All products are available at the beginning of each period. With the customer demands come in, the inventory of each product will be reduced one by one. Here we assume that each customer only places order for one product, no bulk order is allowed. After a certain time, one product will become unavailable say product i. If a customer place order on this product, he will experience stock out. After that, he will turn to another product with new probability $P_{j|L-\{i\}}, j \neq i, j = 0, 1, 2, \ldots, l$. 0 means that the customer will simply leave without any purchasing. With more customer demand arrives, more products will become unavailable. The possibility of experiencing product stock out is increasing. Stock out will always lead to bad customer experience, normally the customers will not try more than 2 times to buy in one place, especially in the online shop. So we assume that if the customer experience stock-out, he/she will only try to buy substitute product once. If that product is also stock-out, the customer will simply leave without purchasing for



(a) Initial Stage



(b) One product is stock out



(c) More than one products are stock out

Figure 5.1: The substitution process at different stage

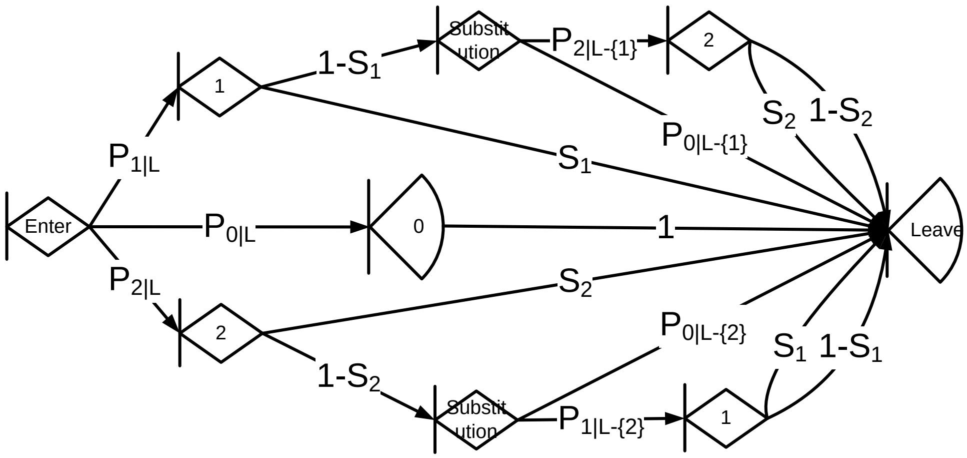


Figure 5.2: The network representation of two products demand substitution

sure. That is $P_{0|L-\{i,j\}} = 1$ where i, j are the products that the customer wants during the purchasing process.

As mentioned above, the major difficulty to model the inventory problem is the stochastic process $\chi_i(t)$. Hopp and Xu (2008) solve the problem by simply replace the product availability process by a constant service rate s_i in the expected effective demand for product i. The constant service rate measures the product availability during the time period $[0, T]$. Their experiments proves that the static approximation performs well. They then model the product substitution process with a flow network. The major assumption they made about the flow network is the "memoryless" property of the network. That means a customer will still try to buy the product even if that product is already stock-out. Smith and Agrawal (2000) also made the similar assumption. But this assumption has a major drawback from accurate modeling. If a customer experiences stock-out when try to buy product i, he/she will never try to choose product i as his/her substitute.

In this research, I will formulate the demand substitution process as a stochastic network and explore the analytical property of this network using GERT technology.

Figure 5.2 shows a network representation of two products substitution. Customer arrives at the shop, he needs to decide which product to purchase or leave the shop without buying. 0 represents leaving without buying. We assume the no-buying option is always part of the product sets. That is $L = \{0, 1, 2\}$. In the first attempt, the customer has the probability $P_{1|L}$ to buy product 1, probability $P_{2|L}$ to buy product 2, and probability $P_{0|L}$ to buy nothing. Note that $P_{0|L} + P_{1|L} + P_{2|L} = 1$. If the customer chooses not buying, he/she will leave with probability 1. If customer chooses to buy product 1, he/she will have the probability s_1 to get the product where s_1 is the service rate of product 1. The customer may also experience stock-out with probability $1 - s_1$. If this case happens, he/she will choose the substitution product. This time product 2 will be chosen with probability $P_{2|L-\{1\}}$, he/she can also choose to leave without buying with probability $P_{0|L-\{1\}}$. If the customer chooses to buy product 2 as a substitution, he/she gets the product with probability s_2 and experience stock-out again with probability $1 - s_2$. Based on our assumption, we will not consider any further substitution. The customer will only try one substitution then leave. The same process can be extended to N products. Figure 5.3 shows the network representation of N products substitution.

Let the profit of each branch as the path-wise additive variable. The network then meets the requirements of the stochastic networks (Philips and Garcia-Diaz, 1981). The profit of selling one product i is $p_i - c_i$. For those branches with probability form like $P_{i|L}$, they are not profit generating branch but the branch simulating the decision process. Thus the profit of that branch is always 0. In most cases, the profit of selling one product is a constant. In some cases, the profit can be a random variable. We will only consider the constant profit case. Thus the W-function is $W_{ij}(v) = p_{ij}M_{ij}(v)$ where p_{ij} is the probability of that branch

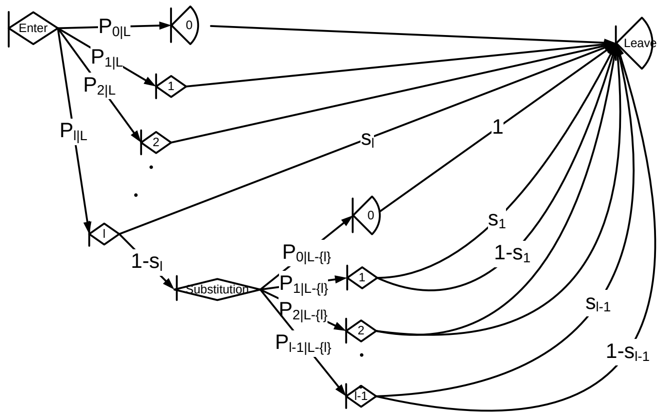


Figure 5.3: The network representation of l products demand substitution

and $M_{ij}(v)$ is the moment generating function of the path-wise additive variable. Because the profit is constant, $M_{ij}(v) = E[e^{v(p_i - c_i)}] = e^{v((p_i - c_i))}$. Furthermore, when the profit of the associated branch is 0, then $M_{ij}(v) = 1$

By have the above formulation, we now have a stochastic network. Then the GERT technology can be applied to answering the following questions:

1. What is the probability that the customer buy a product and leave?

2. What is the probability that the customer leave without buying?

3. What is the mean and variance of the profit of one customer visit?

4. How much profit will be lost if the substitution is not permitted?

Figure 5.4 shows the GERT network with return arc. There is only one first-order loop. The notation of W-function is the node on the path from enter to the next node. For example W_{Els0} is the W-function from the substitution product n with no buying decision. The $0, 1$ after L in the subscript of the W-function stands for the availability, 0 for stock-out and 1 for in-stock. According to Mason's topological equation (Philips and Garcia-Diaz, 1981) with the highest order equal

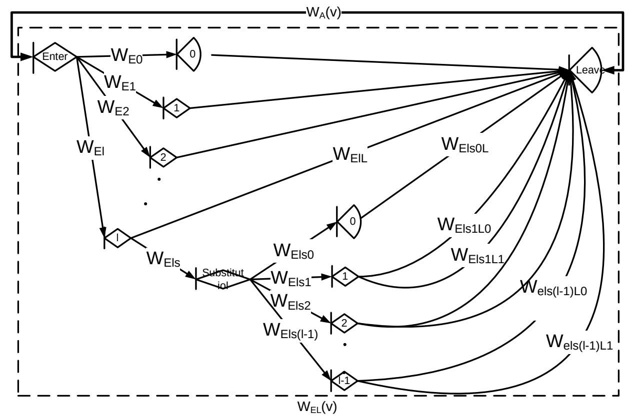


Figure 5.4: The network representation of l products demand substitution with return arc

to 1, we have: $1 - W_A(v)W_{EL}(v) = 0$, that is $W_{EL}(v) = 1/W_A(v)$. The equation can be expended to the following equation (omitting the argument v of the W-functions):

$$W_{E0}W_{E0L}+$$

$$W_{E1}(W_{E1L} + W_{E1s}(W_{E1s0}W_{E1s0L} + W_{E1s2}(W_{E1s2L0} + W_{E1s2L1})+$$

$$\cdots+$$

$$W_{E1sl}(W_{E1slL0} + W_{E1slL1})))+$$

$$\vdots$$

$$W_{El}(W_{ElL} + W_{Els}(W_{Els0}W_{Els0L} + W_{Els1}(W_{Els1L0} + W_{Els1L1})+$$

$$\cdots+$$

$$W_{Els(l-1)}(W_{Els(l-1)L0} + W_{Els(l-1)L1}))) = \frac{1}{W_A} = W_{EL}$$

$$= W_{E0}W_{E0L} + \sum_{i=1}^{l} W_{Ei}(W_{EiL} + W_{Eis}((W_{Eis0}W_{Eis0L}) + \sum_{j=1, j \neq i}^{l} W_{Eisj}(W_{EisjL0} + W_{EisjL1})))$$

Thus the following equation is obtained:

$$W_{EL} = W_{E0}W_{E0L} + \sum_{i=1}^{l} W_{Ei}(W_{EiL} + W_{Eis}((W_{Eis0}W_{Eis0L}) + \sum_{j=1,j\neq i}^{l} W_{Eisj}(W_{EisjL0} + W_{EisjL1}))) \tag{5.4}$$

Now we substitute the W-function with the moment generating function and the probability. Let b_i be the profit of selling product i, d_{ij} be the cost of substituting product i with product j. Note that $b_0 = 0, d_{i0}$ is the shortage cost

$$\begin{aligned}
&W_{Ei} = P_{i|L}e^{v*0} = P_{i|L} && \forall i \in 0, L \\
&W_{EiL} = s_i e^{v*b_i} && s_0 = 1, b_0 = 0 \\
&W_{Eis} = (1 - s_i)e^{v*0} = 1 - s_i && i \neq 0 \\
&W_{Eisj} = P_{j|L-\{i\}} * e^{v*d_{ij}} = P_{j|L-\{i\}}e^{vd_{ij}} && i \neq j, i \neq 0 \\
&W_{EisjL0} = (1 - s_j)e^{v*0} = 1 - s_j && i \neq j \\
&W_{EisjL1} = s_j e^{v*b_j} && i \neq j
\end{aligned}$$

According to Equation 5.1, we have

$$\begin{aligned}
P_{i|L} &= \frac{r_i}{r_0 + \sum_{j\in L} r_j} && \forall i \in \{0, L\} \\
P_{j|L-\{i\}} &= \frac{r_j}{r_0 + \sum_{k\in L} r_k - r_i} && \forall j \in \{0, L\}, i \neq j \\
P_{0|L-\{i,j\}} &= 1 && \forall i, j \in L, i \neq j
\end{aligned}$$

Substitute the above results into Equation 5.4, we have the following equation:

$$W_{EL}(v) = P_{0|L} + \sum_{i \in L} P_{i|L}(s_i e^{vb_i} + (1-s_i) \sum_{j \in \{0,L\}, j \neq i} P_{j|L-\{i\}} e^{vd_{ij}} (1 - s_j + s_j e^{vb_j}))$$

Since the customer will eventually leave for sure, thus $p_{EL} = 1$, $W_{EL}(v) = p_{EL} M_{EL}(v) = M_{EL}(v)$. Thus we have

$$M_{EL}(v) = P_{0|L} + \sum_{i \in L} P_{i|L}(s_i e^{vb_i} + (1-s_i) \sum_{j \in \{0,L\}, j \neq i} P_{j|L-\{i\}} e^{vd_{ij}} (1 - s_j + s_j e^{vb_j})) \tag{5.5}$$

$$\tag{5.6}$$

The expected profit of the l products with substitution is:

$$\begin{aligned} \mu_{EL} &= \left. \frac{\partial M_{EL}(v)}{\partial v} \right|_{v=0} \\ &= \sum_{i \in L} P_{i|L}[s_i b_i + (1-s_i) \sum_{j \in \{0,L\}, j \neq i} P_{j|L-\{i\}}(d_{ij} + s_j b_j)] \end{aligned} \tag{5.7}$$

An immediate observation from Equation 5.7 is that product substitution will improve the expected profit. If product substitution is not permitted, $P_{j|L-\{i\}} = 0$. The second term in Equation 5.7 is 0. Thus the expected profit is less. The same conclusion is reported by Rajaram and Tang (2001). Another immediate observation is that if the inventory of each product is infinite, then the service rate of each product is 1. That is $s_i = 1, i = 1, 2, \cdots, l$. Equation 5.7 becomes $\mu_{EL} = \sum_{i=1}^{l} P_{i|L}(p_i - c_i)$ which is the expected profit of all products.

To calculate the expected profit, we need to calculate the relationship between the service rate and the inventory level. We define the service rate as follows:

For inventory level y, the service rate is defined as $s = f(y) = E[min(y/N), 1]$, where N represents the total number of customer arrivals.

Similar definition of service rate was used by Denechere and Peck (1995) and Hopp and Xu (2008). According to the properties of service rate, function $f(y)$ should have the following characteristic: $f(y)$ is strictly increasing when y is increasing.

The effective demand for product i can be calculated as the direct demand allocated by the product attraction number and the demand to substitute other products. Let I_i denote the effective demand for product i. Then I_i can be calculated as:

$$I_i = N * P_{i|L} + \sum_{j=1, j\neq i}^{l} [N * P_{j|L} * (1 - s_j) P_{i|L-\{j\}}] \tag{5.8}$$

To calculate the function between the service rate and the inventory level, we consider the demand for product i in Figure 5.3. Suppose that N is the total number of customer arrivals in a given period. The number of customer for product i is $NP_{i|L}$. Since product substitution is allowed, product i may be used to substitute the demand for other product. The total number of product i that used to substitute the other products is $\sum_{j=1,j\neq i}^{l} P_{i|L-\{j\}}(1 - s_j)N_j$. Thus the total demand for product i is $NP_{i|L} + \sum_{j=1,j\neq i}^{l} P_{i|L-\{j\}}(1 - s_j)N_j$. According to the definition of service rate given above, we have

$$\begin{aligned} s_i = f(y_i) = E[min\left(\frac{y_i}{NP_{i|L} + \sum_{j=1,j\neq i}^{l} P_{i|L-\{j\}}(1 - s_j)NP_{j|L}}, 1\right)] \\ = E[min\left(\frac{y_i}{N}\frac{1}{P_{i|L} + \sum_{j=1,j\neq i}^{l} P_{i|L-\{j\}}(1 - s_j)P_{j|L}}, 1\right)] \\ = E[min(\frac{y_i}{I_i}, 1)] \end{aligned}$$

Substitute Equation 5.1 into the above equation, we have:

$$s_i = E[min\left(\frac{y_i}{N}\frac{\sum_{i=0}^{l} r_i}{r_i + \sum_{j=1,j\neq i}^{l}\frac{r_i r_j}{\sum_{i=0}^{l} r_i - r_j}(1-s_j)}, 1\right)] \tag{5.9}$$

Equation 5.9 shows the general relationship between the inventory level and the product attraction index. By observing the equation, the following conclusions are obtained:

1. The service rate of all products are affecting each other. Higher service rate requires higher inventory level.

2. Given the same interval level for a product, the service rate will be decreased if the product attraction index is increased because more product will chose that product as substitution and more demand is allocated to that product.

3. By increasing other products service rate, the service rate of this product will also be increased since less products is requesting substitution. This is because the increased inventory reduces the stock-out possibilities of other products which reduces the demand for this product as substitution.

4. Note that if all $s_i = 1$, the effective demand for product i (Equation 5.8 becomes $I_i = N * P_{i|L} = N\frac{r_i}{\sum_{i=0}^{l} r_i}$ which is a standard attraction model.

5.1 Numeric Experiments

We assume that the number of customer arrivals N is normal with mean λ and standard deviation $\sqrt{\lambda}$. We assume $\lambda = 1000$ and consider four scenarios of the product attraction number and the inventory levels. In scenarios 1 and 3, inventory level proportionally match demand for each product. In scenarios 2

Table 5.1: The service rate under different scenarios

	Scenarios	Service Rate
1	$\overrightarrow{r} = (1,1,1,1)\overrightarrow{y} = (250,250,250)$	$0.986, 0.978, 0.985$
2	$\overrightarrow{r} = (1,1,1,1)\overrightarrow{y} = (150,300,450)$	$0.874, 0.999, 1.000$
3	$\overrightarrow{r} = (1,1,2,3)\overrightarrow{y} = (150,300,450)$	$0.997, 0.999, 0.999$
4	$\overrightarrow{r} = (1,1,2,3)\overrightarrow{y} = (250,250,250)$	$1.000, 0.905, 0.857$

Table 5.2: The percentage errors of effective demand%

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Attraction $\overrightarrow{r}$	(1,1,1,1)	(1,1,1,1)	(1,1,2,3)	(1,1,2,3)
Inventory $\overrightarrow{y}$	(250,250,250)	(150,300,450)	(150,300,450)	(250,250,250)
$\kappa = 0.25$	0.73,0.75,0.80	-0.52,0.85,0.51	0.46,0.50,-0.47	0.94,0.83,-0.48
$\kappa = 0.75$	1.00,1.05,1.10	0.98,1.20,-1.21	1.48,1.25,1.91	1.25,-1.15,-1.26
$\kappa = 1$	1.12,1.16,1.12	1.85,1.85,-1.65	1.64,1.53,-1.77	-1.35,1.27,1.48
$\kappa = 4$	3.25,3.29,3.29	3.86,4.46,3.26	4.06,3.40,3.33	3.61,3.38,4.02
$\kappa = 8$	7.00,7.01,7.00	7.49,7.08,8.10	7.91,-7.77,7.51	9.20,8.57,10.22

and 4, inventory levels do not match demand for each product, which results in frequent stock-outs and demand substitution. The profit of selling each product is set to 1.

From Table 5.2, it can be seen that in scenario 1 and 4 the inventory level for each product is the same, but the product attraction index is different. Product 3 has high attraction number in scenario 4, the service rate is thus lower, which satisfy the observation from equation 5.9. In scenario 2, product 1 has less inventory than the average demand. Thus product 1 will experience stock-out frequently. By allowing demand substitution, the service rate is improved.

Table 5.2 shows the percentage error of the effective demand of the stochastic network representation model and the actual demand process model. The experiment result shows that our model is accurate when the variance of customer arrival low. With the increasing variance of customer arrival, the error percentage is also increased.

5.2 Summary

In this research, we considered the demand substitution problem. We model the demand substitution process by stochastic network. The stochastic inventory changing process is substituted by a constant service rate. The stochastic network is then solved by the GERT technology. Our methods significantly reduce the complexity of the dynamic model of demand substitution used by Mahajan and van Ryzin (2001b,a). The experiment shows that the reported relationship between the service rate and inventory level is valid. The experiment also shows that our model is accurate when the variance of customer arrival is low. The error is increased when the variance of customer arrival is increased.

Chapter 6

Conclusions

The overall goal of this work was to study the impact of demand substitution to the supply chain network. In this chapter, section 6.1 summarizes the research conducted in this dissertation. Section 6.2 discusses research contributions and Section 6.3 discusses possible future direction of this research.

6.1 Research Summary

In this research, demand substitution is studied from several perspectives. In chapter 3, I mainly studied the impact of demand substitution on the bullwhip effect. The bullwhip effect is a very important measurement of the supply chain network. A new measurement of the bullwhip effect is proposed to measure not only the fluctuation of amplification but also the frequency of the fluctuation. The research shows that the demand substitution always leads to better profit. The bullwhip effect can be reduced by allowing demand substitution.

In chapter 4, we use the methodology of metamodel to investigate a multi-echelon supply chain network with time based service level agreement. Simulation models are built to study the system behavior and DASE techniques are applied to obtain insights and practical findings.

We look into both single factor metamodel and multi-factors metamodel for a multi-echolon supply chain simulation model. For the single-factor metamodel, the demand inter-arrival rate is set to be the input variable and the average total cost is set to be the output variable. The goal is to reduce the average total cost as much as possible. To calculate the average total cost, the initial inventory and the s and S need to be calculated first. The initial inventory and S are calculated according to the *order-up-to* level of the periodical review model. The value of s is calculated based on the results of continuous review model. The metamodel is tested by applying to the testing samples. The test results show that the metamodel performs well in the range from 0.1 to 1.9. The fitted closed-form expressions for the mean and standard deviation of average total cost are obtained. These expressions can be used as surrogate models to substitute the actual simulation model in its parent model.

For multi-factor metamodel, nine variables are incorporated in the metamodel, DASE analysis shows that the resulting metamodel is significant. Practical policies are inferred from the metamodel. Although this study focuses on the illustrative multi-echelon supply chain problem, we deem that the metamodel methodology, coupled with DASE techniques, can find wide application in many other complicated systems.

We then studied a two echelon supply chain network where the substitution is allowed. The relationship between the inventory and the service rate is obtained.

The results shows that the relationship between service rate and inventory is one-to-one.

In chapter 5, we considered the demand substitution problem. We model the demand substitution process by stochastic network. The stochastic inventory changing process is substituted by a constant service rate. The stochastic network is then solved by the GERT technology. Our methods significantly reduce the complexity of the dynamic model of demand substitution used by Mahajan and van Ryzin (2001b,a). The experiment shows that the accuracy is particularly good for the case of Poisson arrivals and it performs better than the results reported by Hopp and Xu (2008).

6.2 Research Contributions

In this research, the following contributions are made:

1. Mathematically formulated the impact of two products demand substitution to the bullwhip effect under the assumption that a portion of product 1 is always used to substitute product 2. The results shows that the demand substitution can reduce the bullwhip effect of the supply chain network. The bullwhip effect has been widely studied on single product or multi products without demand substitution, little has been reported on the bullwhip effect on multi products with demand substitution. This dissertation studied the bullwhip effect under two products substitution case. The lower limit of the bullwhip effect is obtained and the implied managerial meaning is discussed.

2. Established the input-output relationship of a three echelon supply chain network using Metamodel methodology. The difficulty of stochastic inventory changing process prevents the mathematical modeling of complex supply chain networks. Metamodel methodology is used to obtain the input-output relationship for complex supply chain network.

3. Proposed new stochastic network representation of the demand substitution process. To overcome the difficulty caused by the stochastic inventory change process, the service rate is used to replace the stochastic inventory changing process. Then a stochastic network is constructed to represent the demand substitution process. By solving this network representation, the characteristics of the demand substitution process can be studied. The details are addressed in Chapter 5.

6.3 Future Works

There are many ways to extend the research in this dissertation. First, the impact of the two products substitution to the bullwhip effect can be extended by modeling N products. It can also be extended by changing the demand process. Secondly, we represented the demand substitution process by using stochastic network. It is of great importance. we assume that the product profit is constant in this research and 0 substitution cost. The research can be easily extended to cover these conditions. Thirdly, we use the rational deterministic allocation model to model the customer substitution process.However, literatures from marketing and psychological research suggests that the customer purchase pattern may not be rational at all. The purchase decision is largely influenced by the surrounding environment, social status, emotional condition and other subjective factors. The research of this dissertation can be extended if the demand substitution model could capture the customer behavior more accurately. Finally, in this model we only assume one demand substitution attempt, it might be of interest to model multiple demand substitution attempts.

Bibliography

Bibliography

Agrawal, N. and Smith, S. A. (2003). Optimal retail assortments for substitutable items purchased in sets. *Naval Research Logistics*, 50(7):793–822.

Application Service Provider Industry Consortium (2000). An overviewof the guide to service level agreements. White Paper. `http://www.cisco.com/warp/public/765/emea/aip_asp/documents/SLA_White_Paper.pdf`.

Axsäter, S. (2006). *Inventory Control*. International Series in Operations Research & Management Science. Springer, 233 Spring Street, New York, NY 10013, USA, 2 edition.

Badiru, A. B. and Sieger, D. B. (1998). Neural network as a simulation metamodel in economic analysis of risky projects. *European Journal of Operational Research*, 105(1):130–142.

Barton, R. R. (1992). Metamodels for simulation input-output relations. In *Proceedings of The 24th Conference on Winter Simulation*, pages 289–299, New York, NY, USA. ACM.

Bassok, Yehuda, Anupindi, Ravi, and Akella, Ram (1999). Single-period multiproduct inventory models with substitution. *Operations Research*, 47(4):632–642.

Bitran, G. R. and Dasu, S. (1992). Ordering policies in an environment of stochastic yields and substitutable demands. *Operations Research*, 40(5):999–1017.

Caggiano, K., Jackson, P., Muchstadt, J., and Rappold, J. (2007). Optimizing service parts inventory in a multiechelon, multi-item supply chain with time-based customer service level agreements. *Operations Research*, 55(2):303–318.

Caglar, D., Li, C.-L., and Simchi-Levi, D. (2004). Two-echelon spare parts inventory system subject to a service constraint. *IIE Transactions*, 36(7):655–666.

Chen, F., Drezner, Z., Ryan, J. K., and Simchi-Levi, D. (2000a). Quantifying the Bullwhip Effect in a Simple Supply Chain: The Impact of Forecasting, Lead Times, and Information. *Management Science*, 46(3):436–443.

Chen, F., Ryan, J. K., and Simchi-Levi, D. (2000b). The impact of exponential smoothing forecasts on the bullwhip effect. *Naval Research Logistics*, 47(4):269–286.

Cheng, R. C. H. (1999). Regression metamodeling in simulation using bayesian methods. In Farrington, P. A., Nembhard, H. B., Sturrock, D. T., and Evans, G. W., editors, *Proceedings of the 31st conference on Winter simulation*, pages 330–335, New York, NY, USA. ACM.

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

Craven, B. and Islam, S. (2007). Logistics and supply chain management: a review of some operations research approaches. *International Journal of Industrial and Systems Engineering*, 2(4):444–453.

Denechere, R. and Peck, J. (1995). Competition over price and service rate when demand is stochastic: A strategic analysis. *RAND Journal of Economics*, 26(1):148–162.

Drezner, Z. and Gurnani, H. (2000). Deterministic hierarchical substitution inventory models. *The Journal of the Operational Research Society*, 51(1):129–133.

Drezner, Z., Gurnani, H., and Pasternack, B. A. (1995). An eoq model with substitutions between products. *The Journal of the Operational Research Society*, 46(7):887–891.

Ernst, Ricardo and Kouvelis, Panagiotis (1999). The effects of selling packaged goods on inventory decisions. *Management Science*, 45(8):1142–1155.

Forrester, J. (1958). Industrial dynamics, a major breakthrough for decision makers. *Harvard Business Review*, 36:37–66.

Fuller, J. B., O'Connor, J., and Rawlinson, R. (1993). Tailored logistics: The next advantage. *Harvard Business Review*, 71(3):87–98.

Gilbert, K. (2005). An ARIMA Supply Chain Model. *Management Science*, 51(2):305–310.

Hadley, G. and Whitin, T. M. (1963). *Analysis of Inventory Systems*. Prentice Hall.

Hopp, W. and Spearman, M. (2000). *Factory Physics*. McGraw-Hill Higher Education, second edition.

Hopp, W. J. and Xu, X. (2008). A static approximation for dynamic demand substitution with applications in a competitive market. *Operations Research*, 56(3):630–645.

Ingene, C. and Moinzadeh, K. (1993). An inventory model of immediate and delayed delivery. *Manage. Sci.*, 39(5):536–548.

Kahn, J. A. (1987). Inventories and the volatility of production. *American Economic Review*, 77(4):667–679.

Karaesmen, I. and van Ryzin, G. (2004). Overbooking with substitutable inventory classes. *OPERATIONS RESEARCH*, 52(1):83–104.

Kawagoe, T. and Wada, S. (2005). The Bullwhip Effect: A Counterexample. *Proc. of 2005 IEEE/WIC/ACM IAT*, pages 124–127.

Kleijnen, J. P. (1986). *Statistical tools for simulation practitioners*. Marcel Dekker, Inc., New York, NY, USA.

Kleijnen, J. P. (2007a). Kriging metamodeling in simulation: A review. Discussion Paper 2007-13, Tilburg University, Center for Economic Research.

Kleijnen, J. P. (2007b). Regression models and experimental designs: a tutorial for simulation analysts. In Henderson, S. G., Biller, B., Hsieh, M.-H., Shortle, J., Tew, J. D., and Barton, R. R., editors, *Proceedings of the 39th conference on Winter simulation*, pages 183–194, Piscataway, NJ, USA. IEEE Press.

Kleijnen, J. P. (2008). *Design and Analysis of Simulation Experiments*, volume 111 of *International Series in Operations Research & Management Science*. Springer.

Kleijnen, J. P., Sanchez, S. M., Lucas, T. W., and Cioppa, T. M. (2005). A user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing*, 17(3):263–289.

Kleijnen, J. P. and van Beers, W. C. M. (2005). Robustness of kriging when interpolating in random simulation with heterogeneous variances: Some experiments. *European Journal of Operational Research*, 127(3):826–834.

Law, A. and Kelton, W. (2000). *Simulation Modeling and Analysis*. McGraw-Hill, third edition.

Lee, H. L., Padmanabhan, V., and Whang, S. (1997a). The bullwhip effect in supply chains. *MIT Sloan Management Review*, 38(38):93–102.

Lee, H. L., Padmanabhan, V., and Whang, S. (1997b). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 43(4):546–558.

Lee, H. L., Padmanabhan, V., and Whang, S. (2004). Information distortion in a supply chain: The bullwhip effect. *Management Science*, 50:1875–1886.

Leemis, L. M. (2004). Building credible input models. In Ingalls, R. G., Rossetti, M. D., Smith, J. S., and Peters, B. A., editors, *Proceedings of the 36th conference on Winter simulation*, pages 29–40. Winter Simulation Conference.

Leng, M. and Palar, M. (2008). Allocation of cost savings in a three-level supply chain with demand information sharing: a cooperative-game approach. *Operations Research*, 56(6).

Mahajan, S. and van Ryzin, G. (2001a). Inventory competition under dynamic consumer choice. *Operations Research*, 49(5):646–657.

Mahajan, S. and van Ryzin, G. (2001b). Stocking retail assortments under dynamic consumer substitution. *Operations Research*, 49(3):334–351.

McGillivray, A. R. and Silver, E. A. (1978). Some concepts for inventory control under substitutable demand. *INFOR*, 16(1):47–63.

Metters, R. (1997). Quantifying the bullwhip effect in supply chains. *Journal of Operations Management*, 15(2):89–100.

Milner, J. M. and Olsen, T. L. (2008). Service-Level Agreements in Call Centers: Perils and Prescriptions. *Management Science*, 54(2):238–252.

Netessine, S. and Rudi, N. (2003). Centralized and competitive inventory models with demand substitution. *Operations Research*, 51(2):329–335.

Parasuraman, A., Zeithaml, Valarie A., and Berry, Leonard L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49(4):41–50.

Parlar, M. (1988). Game theoretic analysis of the substitutable product inventory problem with random demand. *Naval Research Logistics*, 35(3):397–409.

Parlar, M. and Goyal, S. (1984). Optimal ordering decisions for two substitutable products with stochastic demands. *Opsearch*, 21:1–15.

Philips, D. T. and Garcia-Diaz, A. (1981). *Fundamentals of Network Analysis*. International series in industrial and system engineering. Prentice-Hall, hardcover edition.

Rajaram, K. and Tang, C. S. (2001). The impact of product substitution on retail merchandising. *European Journal of Operational Research*, 135(3):582–601.

Reis dos Santos, M. I. and Porta Nova, A. M. (2006). Statistical fitting and validation of non-linear simulation metamodels: A case study. *European Journal of Operational Research*, 127(1):53–63.

Robinson, L. W. (1995). Optimal and approximate control policies for airline booking with sequential nonmonotonic fare classes. *Operations Research*, 43(2):252–263.

Ryan, J. K. (1997). *Analysis of Inventory Models with Limited Demand Information*. Doctor of philosophy, Northwestern University, Evanston, Illinois.

Sahin, F. and Robinson, E. P. (2006). *Handbook of Global Logistics and Supply Chain Management*, chapter Inventory Management in Global Supply Chains, page 195. Sage Publications.

Santos, I. R. and Santos, P. R. (2007). Simulation metamodels for modeling ouput distribution parameters. In Henderson, S. G., Biller, B., Hsieh, M.-H., Shortle, J., Tew, J. D., and Barton, R. R., editors, *Proceedings of the 2007 Winter Simulation Conference*, pages 910–918.

Silver, E. A., Pyke, D. F., and Peterson, R. (1998). *Inventory management and Production Planning and Scheduling*. John Wiley & Sons, 3 edition. ISBN-10: 0471119474.

Smith, S. A. and Agrawal, N. (2000). Management of multi-item retail inventory system with demand substitution. *Operations Research*, 48(1):50–64.

Song, L., Li, X., and Garcia-Diaz, A. (2008). The influence of demand substitution to a two echelon supply chain. In Fowler, J. and Mason, S., editors, *Proceedings of the 2008 Industrial Engineering Research Conference*.

Spearman, M. and Hopp, W. (2001). *Factory Physics Second Edition*. McGraw Hill Higher Education.

Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35(3):321–339.

Tijms, H. C. (1972). *Analysis of (s,S) inventory models*. Amsterdam, Mathematisch Centrum.

Wagner, H. M. and Whitin, T. M. (1958). Dynamic version of the economic lot size model. *Management Science*, 5(1):89–96.

Wong, H., Kranenburg, B., van Houtum, G.-J., and Cattrysse, D. (2007). Efficient heuristics for two-echelon spare parts inventory systems with an aggregate mean waiting time constraint per local warehouse. *OR Spectrum*, 29(4):699–722.

Wong, W., Jaruphongsa, W., and Lee, L. H. (2008). Supply chain performance measurement system: a monte carlo dea-based approach. *International Journal of Industrial and Systems Engineering*, 3(2):162–188.

Zipkin, P. H. (2000). *Foundations Of Inventory Management*. Irwin Professional Pub.

Vita

Laigang Song received a B.S. degree in Metallurgy Engineering from Northeastern University, CHINA in 2000 and a M.S. degree in Mechanical Engineering from Chinese Academy of Science in 2003. He won several prizes for academic excellence during his undergraduate program. He won Liu Yongling scholarship during his master program. After his master, he worked as a software engineer for Shenyang Institute of Automation Research, Clausmark Consulting GmbH between 2003 and 2005. He started Ph.D. study in Industrial Engineering at the University of Tennessee in 2006. He is the feature star graduate student and won the most outstanding Ph.D. student award in 2009. During his Ph.D. program, he taught an undergraduate engineering simulation course with full responsibilities for one semester. He was teaching assistant for several other courses. He was involved in various research projects in the department of industrial and information engineering. He was also in charge of the department web site, ftp site, version control server. Song has published or submitted 8 articles for publication in journals and conferences. His areas of research include operations research, data mining, simulation and modeling, supply chain management, signal and image processing. Song is a member of several professional bodies including the Institute of Industrial Engineers (IIE), American Society for Quality (ASQ), Institute for Operations Research and the Management Sciences (INFORMS). He can be reached at songlaigang@gmail.com. He recently accepted a position at INEX technology, Knoxville TN. Song is married to his college sweetheart, beautiful Dongchun Qiao; they both have one lovely daughter Selina Jiayin Song.