IDENTIFICATION AND MITIGATION OF BOTTLENECKS IN COMPLEX PRODUCTION SYSTEM

A Thesis by

Pradeesh Narayanasamy

Bachelor of Engineering, Anna University, India, 2007

Submitted to the Department of Industrial and Manufacturing Engineering
and the Faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Master of Science

December 2010

IDENTIFICATION AND MITIGATION OF BOTTLENECKS IN COMPLEX PRODUCTION SYSTEM

The following faculty members have examined the final copy of this thesis for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Master of Science with a major in Industrial Engineering.

Krishna K. Krishnan, Committee Chair

Mehmet B. Yildirim, Committee Member

Ramazan Asmatulu, Committee Member

ACKNOWLEDGMENTS

It is a great pleasure to thank all the people who made this thesis possible.

I owe my deepest gratitude to my advisor Dr. Krishna K. Krishnan who helped me in almost all the situations I encountered at Wichita State University. I developed a habit of reasoning and validating each and every step in my research and life from him.

I thank my committee members Dr. Bayram Yildirim and Dr. Ramazan Asmatulu who took time to review my thesis and gave some valuable suggestions which made my work better.

I also thank the members of Facilities Planning and Logistics Research Group who helped me in the completion of my thesis.

I would like to specially thank Prasanna, Prakash, Sathya, Thiagaraj, Mervyn, Govind, Vishnu and Ravinder for their valuable time to discuss and give suggestions invaluable to my research and constant support during my time of need.

Finally, I would like to dedicate this thesis to my grandma, parents, and brother who are the most precious people in my life.

ABSTRACT

A production system is a system where the raw materials are converted to finished products. A production system is classified in to two types based on the products processed namely single product production system and multi-product production system. The products in any production system would be in either value added or non-value added state. This research focuses on reducing the non-value added state in multi-product or complex production system by analyzing the bottlenecks. A bottleneck machine causes blocking or starving of parts in the system thereby increasing the non-value added time and reducing the system performance. Bottlenecks can be mitigated by control strategies such as buffer allocation and capacity addition. In order to mitigate the bottlenecks, the location, source and type of bottleneck in any system has to be identified. This research uses multiple metrics in order to identify the bottleneck and its type. Based on the metric values, the control strategies are implemented by the developed heuristics such as buffer allocation based on qualitative characteristics, capacity addition based on highest utilization and economic analysis based on sensitivity analysis. Multiple options are given with respect to their performance improvement for the management or the customer to select in order to give flexibility in terms of investment, demand and layout space.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Production System

A production system processes a set of raw materials in to finished goods based on the requirement of the customer. Production systems can be classified into many types based on the type of layout, products processed, volume of products etc. One of the problems faced by any production system is bottleneck machines caused by the failures and variability in the system. These bottlenecks should be handled for a smooth flow of products and for an optimal system performance. This research focuses on such problem dealing with bottleneck mitigation in multi-product production system.

1.2 Facility Layout

The facility layout design plays a vital role in improving the performance of a system. The layout is designed based on the type of product or process. The main objectives for a layout design are the proper allocation of machines and other resources based on the material handling costs and throughput. The expenses associated with operations can be improved up to 30% by effective facility layout design (Tompkins and White, 1996). The previous researchers have developed methodologies for arrangement of machines considering the material handling costs and throughput.

1.3 Production Systems

A production system consists of set of machines in which raw materials are processed into a finished product. The products in any system will either be in value added activity such as processing or non-value added state such as waiting in a buffer or machine for processing. Process improvement activities attempt to reduce non-value added time. One of the reasons for increased non-value added time in a production system is due to bottleneck machine(s). Previous researchers have defined and identified these bottlenecks using different methods. But most of their research focused on single product production systems.

Due to globalization, industries are required to manufacture products with different specifications and design to meet the increasing competition and customer demand. Thus, most of the present industries such as automotive and electronics have switched to mixed model production. Therefore, there is a necessity for developing new methodologies for improving the performance based on multi-product systems. Bottlenecks in multi-product production systems are not well defined because of the complexity of the problem. This research focuses on handling bottlenecks using buffers. Previously, very few researchers considered complex production systems for allocating buffers in a production system.

In this research, different types of multi-product production systems are considered for defining a new methodology in buffer allocation. Most of the present day production systems are customized for the specific company‟s requirement and so some systems are well balanced. One of the major drawbacks in a balanced system is that a small failure could affect the whole system. If there is a problem in one machine, then it blocks the flow of products throughout the system. This problem can be overcome by placing buffers in between machines. This helps in maintaining a smooth flow of products even in the presence of variability. The different systems considered in this research cover most of the present industrial scenarios for buffer allocation.

- Single line multi product production systems or sequential flow
- Products with alternate and multiple routes or jumbled flow

1.4 Bottlenecks in Production System

The flow of products in any system is disrupted due to machine failure and operator failure in the system. If these failures occur repetitively, then the machines causing these failures are bottlenecks. The failure in a single machine would disrupt the whole system. So, the important process improvement technique is to mitigate the bottlenecks. Bottleneck identification is the first step to mitigate the bottlenecks. There are many researches for identifying bottleneck machine such as longest queue length (Lawrence & Buss, 1994), lowest production rate (Kuo, Lim, & Meerkov, 1996), Buffer with high work in process, (Kuo, Lim, & Meerkov, 1996), Highest ratio of sensitivity, (Kuo, Lim, & Meerkov, 1996), Lowest blockage and starvation time, (Chiang, Kuo, & Meerkov, 1998) and (Li, Chang, & Ni, 2008), Highest utilization, (Law, & Kelton, 2000), Longest active duration, (Roser, Nakano, & Tanaka, 2002).

1.5 Role of Buffer in Production System

Throughput is the most significant performance parameter in a production system. The throughput is affected by the variability in a system such as processing time variability and machine failure rate. The variability in a machine „m‟either causes blocking and starving for the next and previous machine. Blocking occurs if there is an over production in the machine „m-1‟ or if there is a breakdown at the „m+1‟ station. Starving occurs if the machine „m-1‟ breaks down or if the processing time of the „m+1‟ station is very low. If these scenarios occur in a well balanced line, the whole system would be affected. So the bottleneck stations have to be improved in order to avoid blocking or starving. The variability can be handled by some expensive methods such as, adding parallel machines and process design change, and layout change. One of the alternate methods to handle variability is by adding buffers.

Buffers are placed in between machines in order to handle blocking or starving. If there is a breakdown in the „m+1‟ station, buffers would hold the part to avoid blocking and vice-versa. These scenarios are very common in industries such as automobile, printer assembly, etc. As the variability increases, buffer size also increases. As a result of increase in buffer size, the cost involved in implementation and space provided also increases, so an optimal size of buffer should be placed before machines for the required throughput. According to Conway, Maxwell, McClain, and Thomas (1988), work allocation and buffer size are the two significant factors affecting throughput. The optimal number of buffers to be placed in a system is a result of balancing between the implementation cost and production losses (Faria, Matos, and Nunes 2006). The optimal location of buffers also plays an important role in a functional layout as a variety of products are processed in the same layout. The unbalanced lines with processing time variability perform better than the balanced lines, if buffers are placed (Patti, Watson, and Blackstone, 2008). Previous researchers have developed methodologies for an optimal buffer allocation in different production scenarios.

1.5.1 Single Product Production Systems

Hillier and Boling (1966) observed that the performance of transfer lines were better when more buffer units are arranged to the two ends and the phenomenon is called „bowl‟ phenomenon. Anderson and Moodie (1969) developed mathematical models and solved buffer capacity for steady-state production line systems using simulations. Powell (1994) analyzed a three station unbalanced asynchronous line and indicated that the buffers should be allocated to bottleneck stations and size should increase towards the center of the line. Also, he demonstrated that the mean imbalance is more sensitive for buffer allocation than variance imbalance. Bulgak, Diwan and Inozu (1995) used genetic algorithm to solve the optimal buffer size in asynchronous assembly systems. Powell and Pyke (1996) examined factors such as line length, multiple bottlenecks and coefficient of variation with buffer allocation. Monotonically increased buffers along the downstream showed better performance for both balanced and unbalanced lines (So, 1997). Powell, and Pyke (1998) extended their previous work by developing heuristics for unbalanced assembly systems and found both mean and variance should be considered for buffer size determination. Gardiner and Blackstone (1998) analyzed Goldratt"s dynamic buffering model by reducing the level of work in process inventory and improving the output of the system by protecting non-constraints.

Hillier (2000) studied production lines with variable processing times following exponential and Erlang distribution. Even though the inverted bowl phenomenon is optimal for buffer allocation in many cases, more buffers are needed near the bottleneck stations for an optimal solution. Gershwin and Schor (2000) considered continuity, monotonicity, and concavity as the qualitative properties to solve the primal-dual problem by gradient method for minimizing buffer space and maximizing production rate. Papadopoulos and Vidalis (2001) found an optimal buffer size in an exponentially unreliable and unbalanced six machine production line using an algorithm based on sectioning approach to increase its throughput. Chiadamrong and Limpasontiong (2003) studied the relationship between with bottleneck stations and buffer factors in unbalanced lines. Shi and Men (2003) developed a hybrid algorithm incorporating Tabu Search heuristic in to the Nested Partitions framework. Nahas, Ait-Kadi, and Nourelfath (2006) applied a degraded ceiling approach which gave better results over simulated annealing approach in an unreliable machine environment. Hillier and Hillier (2006) used a cost based model for a simultaneous optimization for work and buffer allocation for each machine in unpaced production lines. Sabuncuoglu, Erel, and Gocgun (2006) studied single and multiple bottleneck stations for buffer allocation to maximize throughput by a heuristic procedure.

Faria, Matos, and Nunes (2006) explained the need for work-in-process buffers of a just-in-time manufacturing system to improve the reliability of a production system. Riberio, Silveria, and Qassim (2007) proposed a mixed integer linear programming model for improving the utilization of a capacity constrained resource by joint optimizing its maintenance and inlet buffer size. Qudeiri, Yamamoto, Ramli, and Al-Momani (2007) proposed a production simulator based on a genetic algorithm and a discrete event (simulation/optimization)for determining an optimal buffer size and minimizing the required number of generations needed to determine a buffer size for a complex production system. Othman, Kamaruddin, and Ismail (2007) discussed on optimal buffer allocation for short unpaced production line based on the shape of mean processing time of that line. Further, Nahas, Ait-Kadi, and Nourelfath (2009) formulated a design on buffer and parallel machine allocation in unreliable production lines to increase its throughput.

Most of the researchers considered either process time variability or machine failure rate in balanced and unbalanced lines and some researchers contributed their work in serial production lines.

1.5.2 Multi-Product Production System

In the growing competition, production companies need to expand their range of products with the available resources which paved the way for multi-product production system. Even though, there are many researches in a single product system, none of the researchers focused on bottleneck detection in multi product system and very few addressed buffer allocation in a multi-product production systems. Nieuwenhuyse, Vandaele, Rajaram, and Karmarkar (2007) developed a queuing model for determining the buffer size for the required service level in a multi product multi reactor batch processing environment. Ye and Han (2006) developed a method for optimal size for stock buffer in front of a bottleneck station in a multi product production environment.

In this research, bottlenecks are identified using longest queue length method and analyzed using developed metrics. Methods for mitigating bottlenecks by buffer allocation and capacity allocation in multi product systems is developed. Mathematical model for an optimal buffer determination is more complex and might have limitation to specific type of systems. Therefore, a generic method is to be developed to address most of the manufacturing scenarios.

1.6 Objective

Based on the literature it is clear that there is a need for a generic method for buffer allocation in complex production systems. So the objective for this research is determined as follows.

- To develop a metric based method to identify bottlenecks in complex production
- To develop heuristics to mitigate bottlenecks in complex production systems by allocating buffer and additional capacity
- To find the economic feasibility of buffer allocation and capacity addition

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

A detailed literature review is listed in different sections of this chapter. Effect of bottlenecks and bottleneck identification is explained in section 2.2. Section 2.3 discusses about various bottleneck mitigation methods. There are many researches in allocating buffer in single product systems which is discussed in section 2.4. Section 2.5 reviews the literature on methods of buffer allocation. Section 2.6 explains the need for an economic buffer rather than a buffer size for maximum throughput.

2.2 Bottlenecks in Production System

According to Boysen, Fliedner, & Scholl (2007), a production line processes a raw-material and it is converted into a finished product after a set of value added activities. There are different types of production system and they are classified based on the types of product processed, processing time between stations, constraints adopted etc. There is a need for process improvement in any industry to remain competitive in the market. One of the problems faced by any production industry is disruption of the work flow by various failures. These failures in any machine cause machines in upstream or downstream to starve or block the products flowing through the system.

The flow of products in any system is disrupted due to various failures such as machine failure, operator failure, power failure and material failure. If these failures occur repetitively, then the machines causing these failures are bottlenecks. The failure in any machine would disrupt the whole system. So, one of the important process improvement technique is to mitigate the bottlenecks. There is a 30 to 40 % reduction in the system efficiency due to bottleneck machines (Chiang, Kuo, & Meerkov, 2001). Bottleneck identification is the first step to mitigate the bottlenecks.

There are many researches for identifying bottleneck machine as follows. Lawrence and Buss, (1994) proposed the longest queue length method where the machine with the longest queue of products becomes a bottleneck. Kuo, Lim, and Meerkov (1996) proposed that the machine which has the lowest production rate becomes a bottleneck as it reduces the flow of products in the whole system. Chiang, Kuo, and Meerkov (1998) proposed that the machine with lowest blocking and starving time becomes a bottleneck. Similarly, the machine with highest utilization becomes a bottleneck machine (Law and Kelton, 2000). Roser, Nakano, and Tanaka (2002) proposed that the machine with the longest active duration as a bottleneck machine. All the previous research focused on bottleneck identification in single product production systems. None of the previous research considered multi product production systems because of the complexity of the problem.

Tamilselvan, Krishnan, and Cheraghi (2010) developed new metrics for identifying bottlenecks based on the inactive duration method. The metrics developed are bottleneck time ratio, bottleneck ratio, bottleneck shifting frequency, and bottleneck severity ratio. Since they considered single product production systems, there is a need for developing new metrics based on multi-product systems to identify the bottleneck machines. Since all these methods and metrics were used in single product production systems, their validity for complex production systems were tested using sample case studies.

Once the bottlenecks are identified, methods have to be developed for mitigating it. Previously, Chiadamrong, and Limpasontipong (2003) proposed that allocating buffer to bottleneck machine would solve blocking or starving there by mitigating bottleneck machines.

Tamilselvan, Krishnan, and Cheraghi (2010) used additional capacity and buffers to mitigate bottlenecks. So, methods are needed to allocate buffer and machines for mitigating bottlenecks in multi-product production system.

2.3 Bottleneck Mitigation

As discussed earlier, temporary bottlenecks caused by blocking or starving can be mitigated by allocating buffers and machines to improve the performance of the system such as throughput. Throughput is the most significant performance parameter in a production system. Throughput which is affected by variability has to be handled in order to have an optimized output. Since adding additional capacity is not necessary in many cases as the bottleneck might be existing for a very short time. Allocating buffers mitigates or eliminates short time bottlenecks. As the variability increases, buffer size also increases. As a result of increase in buffer size, the cost involved in implementation and space provided also increases, so an optimal size of buffer should be placed before machines for the required throughput.

According to Conway, Maxwell, McClain, and Thomas (1988), work allocation and buffer size are the two significant factors affecting throughput. Their study included lines with and without buffer, balanced lines with processing time variability, unbalanced lines and unreliable production lines. The conclusions from their research are as follows;

- Buffer placement in the center is more significant than end allocations
- The production losses is mainly due to the degree of variability rather than the length of line
- The improvement in throughput decreases with buffer increase
- Buffers near bottleneck stations are more significant

The optimal number of buffers to be placed in a system is a result of balancing between the implementation cost and production losses (Faria, Matos, and Nunes 2006). In this study, the reliability for the production systems is considered based on different design issues such as machine redundancy, layout and the maintenance policies. The optimal location of buffers also plays an important role in a functional layout as a variety of products are processed in the same layout. The unbalanced lines with statistical fluctuations perform better than the balanced lines, if buffers are placed (Patti, Watson, and Blackstone, 2008). In their study, Kanban and drum-buffer-rope are compared and the behavior pattern between them was analyzed. There are many researches with different methodologies for optimal buffer allocation in different production scenarios which are explained in the following sections.

2.4 Buffer Allocation in Single Product Systems

Hillier and Boling (1966) observed that the performance of transfer lines were better when more buffer units are arranged to the two ends of a transfer line and the phenomenon is called „bowl‟ phenomenon. Bowl phenomenon was later proved and disproved with different types of systems by other researchers.

Anderson and Moodie (1969) developed mathematical models and solved buffer capacity for steady-state production line systems using simulations. They studied two steady state models and one non-steady state model. They also analyzed the possibility of using time – based buffer as the production lines operate at a slower output rate at the initial transient stage.

Schragenheim and Ronen (1990), analyzed a drum-buffer-rope production and recommended three times the lead time of bottleneck as an optimal buffer size. This conclusion is not applicable to all the production lines since bottleneck might shift from one machine to another and the buffers itself would not improve the performance of the line if the lead time of the bottlenecks are too long.

Hillier, So, and Boling (1993) discussed about the optimal buffer allocation in balanced lines with identical exponential processing time. Their findings are that the amount of buffer spaces plays a major role than the pattern of any buffer allocation in terms of throughput and the buffers tend to be in center of stations for maximum throughout in most cases.

Bulgak, Diwan, and Inozu (1995) used genetic algorithm to solve the optimal buffer size in asynchronous assembly systems. Their study considered an automated manufacturing for allocating buffer to its material handling system. This system needed a tradeoff between a large buffer quantity and a small quantity as large quantities increased the in-process inventories and the small quantity blocked the whole line due to insufficient quantity. Even though their system had stochastic fluctuations, the genetic algorithm gave solutions with reasonable accuracy.

Powell and Pyke (1998) analyzed a three station unbalanced asynchronous line and indicated that the buffers should be allocated to bottleneck stations and size should increase towards the center of the line. Also, it demonstrated that the mean imbalance is more sensitive for buffer allocation than variance imbalance. They also examined factors such as line length, multiple bottlenecks and coefficient of variation with buffer allocation. They addressed issues on allocating limited buffers in unbalanced assembly systems and developed a heuristic based approach to improve the conditions of the systems. They also verified the significance of Alternation Principle and found that this principle works unless the systems are significantly unbalanced. Since there are problems in allocating buffer between high mean station and stations with high variance, the quadrant method is explained and an example is shown in Figure 2.3.

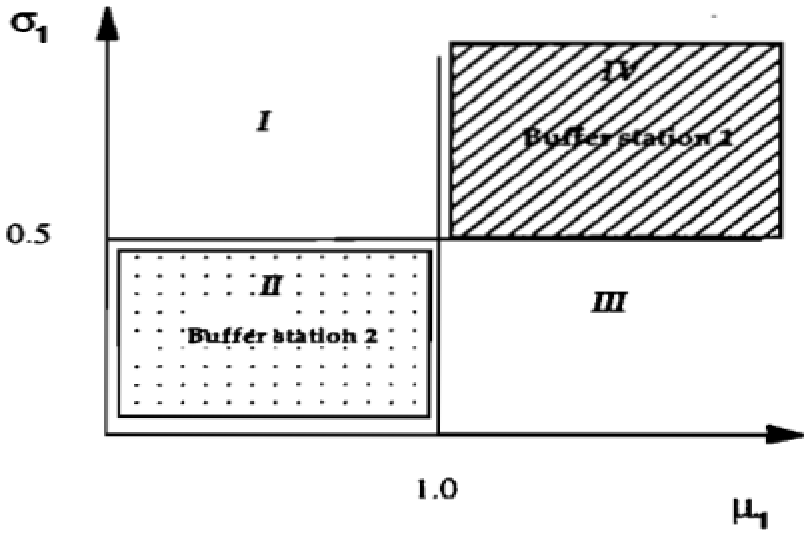

Figure 2-1 Quadrant representations for buffer allocation (Powell and Pyke, 1996)

Monotonically increased buffers along the downstream showed better performance for both balanced and unbalanced lines (So, 1997). They also suggested allotting the slowest processing jobs at the beginning of the line to minimize the work in process inventory. Powell and Pyke (1998) extended their previous work by developing heuristics for unbalanced assembly systems and found both mean and variance should be considered for buffer size determination. Asynchronous assembly lines were considered in this research where managers were facing problems such as rapid product change. Heuristics were developed for buffer allocation for improving the system performance.

Radovilsky (1998) developed a quantitative approach for determining the time buffer required in order to mitigate the disruptions involved in a production system. In order to increase the profit, the machines which are capacity constrained could not be idle and the time buffer is allotted to the system for a continuous supply of parts. Yamashita and Altiok (1998) developed a dynamic programming based on the decomposition type approximation method for production system with phased processing time.

Lutz, Davis, and Sun (1998) developed a simulation and tabu search based heuristic procedure for allocating buffer and deciding the location for buffer. Serial balanced production lines are considered to validate the method. This method has two different routines namely swap routine swap routine and global search routine for identifying the optimal buffer levels with maximum throughput and the location of buffers in the system where buffers are optimal. In the results, the patterns found by previous authors such as inverted bowl phenomenon and symmetrical storage pattern were optimal in most of the cases. It is also arrived that the bottleneck stations need more buffers when compared to other stations and the need for buffers decreases if the degree of imbalance increases.

Gardiner and Blackstone (1998) allotted time buffer to prevent from non constraint resources to produce many parts. If there is a large output from the non constraint resources then the total work in process (WIP) inventory increases in the system which decreases the total system performance. So, we need to maintain the WIP by controlling the output rate of the non constraint machine and the buffer after that machine. In this research, they concluded that dynamic buffering was more efficient than fixed buffering. The system exhibits less variability at initial stages which needs fewer buffers thereby decreasing the work in process inventory.

Harris and Powell (1999) allotted an optimal buffer for reliable unbalanced production lines using a simple search algorithm. They also used simulation to estimate the throughput. The simulation run length is determined based on the precision requirement in the system. The initial systems considered for validating the model were with known optimal buffer allocation and it was further extended with unknown cases. The algorithm gave optimal solutions in balanced lines and near optimal in most of the unbalanced lines.

Hillier (2000) studied production lines with variable processing times following exponential and Erlang distribution. Even though the inverted bowl phenomenon is optimal for buffer allocation in many cases, more buffers are needed near the bottleneck stations for an optimal solution. Their research used a cost model since there was a constraint in the total buffer space. The inverted bowl phenomenon was optimal for balanced lines but it became more pronounced if here was a surge in buffer sizes. Also the bowl pattern diminished for unbalanced lines as the buffer size increased near the bottleneck station. Buffer allocation to the interior stations has to be considered carefully since the disruption of interior machines would affect both the downstream and upstream machines of the system. The buffer size was proportional to the coefficient of variability.

Gershwin and Schor (2000) developed and described various algorithms for buffer space allocation. They also considered continuity, monotonicity, and concavity as the qualitative properties to solve the primal-dual problem by gradient method for minimizing buffer space and maximizing production rate. The developed algorithm was implemented in various systems and can be further extended to multi-product systems with some modifications in the modeling method.

Jeong and Kim (2000) proposed a solution procedure to design the optimal layout for a desired throughput and cost. The design factors included capacity of machines and capacity of buffers. Three different heuristics were proposed based on the output rate or efficiency of the machines. Once the machines are allocated, the buffer allocation solution is found by a local search heuristic.

Papadopoulos and Vidalis (2001) found an optimal buffer size in an exponentially unreliable and unbalanced six machine production line using an algorithm based on sectioning approach to increase its throughput. Huang, Chang, and Chou (2002) proposed a new method for buffer allocation in a flow-shop-type production system. This method used dynamic programming for analyzing system performance measures and thereby allocating buffer. The system has some buffer space constraints and there is a finite capacity in each buffer location before machines within which the buffer size has to be determined in order to improve performance measures such as throughput, minimize work in process inventory and blocking or starving time of the products in the system.

Chiadamrong and Limpasontiong (2003) studied the relationship between with bottleneck stations and buffer factors in unbalanced lines. According to the authors, a bottleneck station holds down the whole system‟s capacity. The source of variability considered in their study was difference in processing time between stations. Since the buffers solved the problem of blocking and starving and there is not enough study regarding relationship between bottleneck machines and buffer characteristics, their study concentrated on those two factors. The analysis of variance considered various factors and there were significant factors such as bottleneck position, mean processing time, variance of processing time, location and size of buffers and the interaction between them. The most important location for a buffer is near the bottleneck machine and it yields the best performance when compared to other allocations.

Shi and Men (2003) developed a hybrid algorithm incorporating Tabu Search heuristic in to the Nested Partitions framework for optimal allocation of buffers. Nahas, Ait-Kadi, and Nourelfath (2006) applied a degraded ceiling approach which gave better results over simulated annealing approach which is shown in Figure 2.4. In their research, the objective was to determine near optimal buffer solution in order to maximize the throughput in large production lines. The total number of buffer spaces is constrained and the allocation should be within that limit. In this approach, there is a tradeoff between the quality of buffer allocation and the search time.

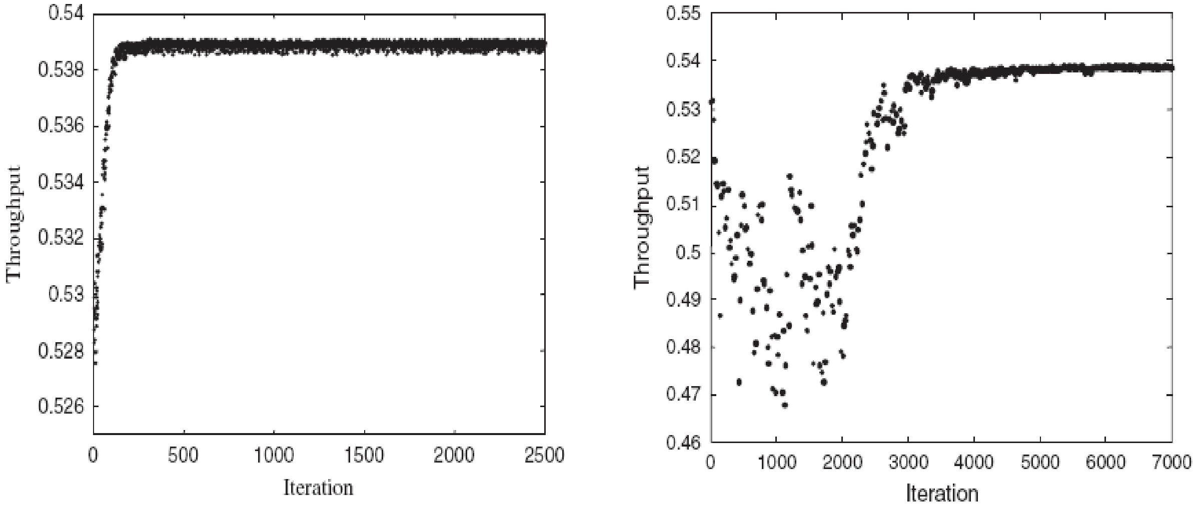

Figure 2-2 Degraded ceiling approach vs. simulated annealing (Nahas et al., 2006)

Sabuncuoglu, Erel, and Gocgun (2006) proposed a new method to allocate buffer in between stations in both reliable and unreliable stations. One of the two objectives is to characterize the optimal buffer allocation. The first step of buffer allocation is to transfer the lowest utilized buffer space to the highest utilized buffer space and the algorithm is shown in Figure 2.3. The buffer space before the bottleneck machine and the buffer space after the bottleneck machine attract more buffers. The buffers are allocated before a bottleneck station which is identified by average production rate method (R_{avg}).

Hillier and Hillier (2006) used a cost based model for a simultaneous optimization for work and buffer allocation for each machine in unpaced production lines. Most of the previous research concentrated in work allocation and buffer allocation separately. In this research, the cost model developed is basically a tradeoff between the profit per unit of throughput and the cost of allotting a buffer space. The interaction between bowl phenomenon on both work load and buffer allocation is investigated. One of the simultaneous allocations is shown in Figure 2.4 below.

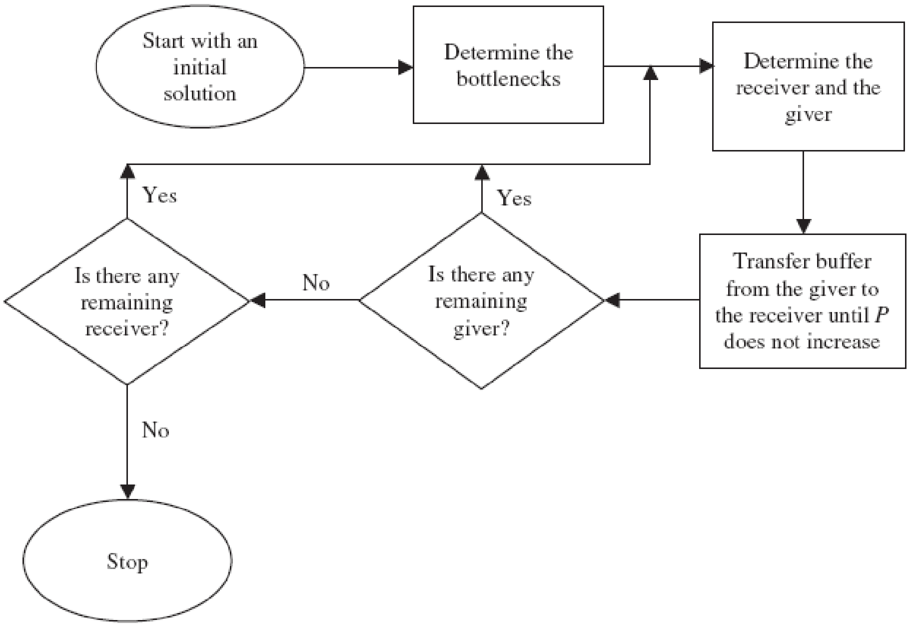

Figure 2-3 Flowchart for algorithm (Sabuncuoglu, Erel, and Gocgun, 2006)

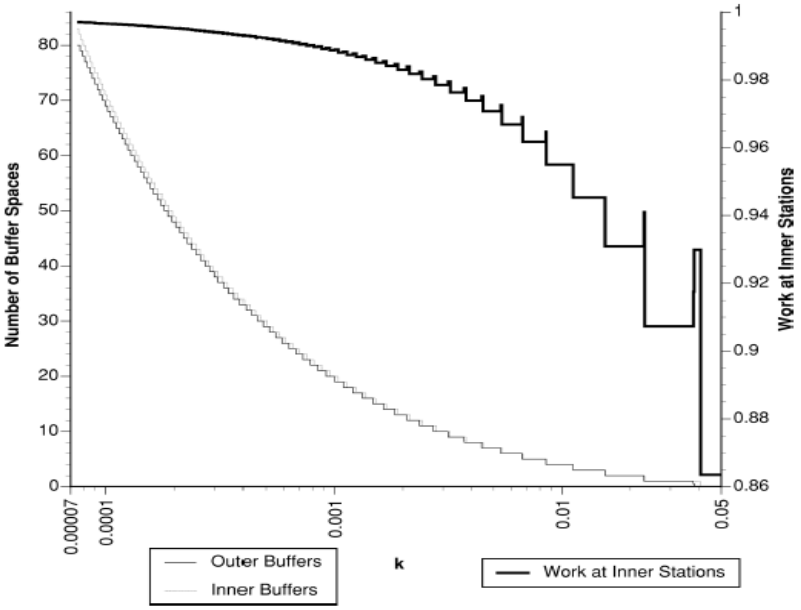

Figure 2-4 Optimal workload and buffer allocations (Hillier and Hillier, 2006)

Erel and Gocgun (2006) studied single and multiple bottleneck stations for buffer allocation to maximize throughput by a heuristic procedure. Faria, Matos, and Nunes (2006) explained the need for work-in-process buffers of a just-in-time manufacturing system to improve the reliability of a production system. Buffers are needed to handle both endogenous and exogenous failures. The research work deals with several design issues in a production line such as maintenance policies, cell layouts, and equipment redundancy. The procedure is explained using a case study with Just in Time manufacturing.

Riberio, Silveria, and Qassim (2007) proposed a mixed integer linear programming model for improving the utilization of a capacity constrained resource by joint optimizing its maintenance and inlet buffer size. Qudeiri, Yamamoto, Ramli, and Al-Momani (2007) proposed a production simulator based on a genetic algorithm and a discrete event (simulation/optimization) for determining an optimal buffer size and minimizing the required number of generations needed to determine a buffer size for a complex production system. Othman, Kamaruddin, and Ismail (2007) discussed on optimal buffer allocation for short unpaced production line based on the shape of mean processing time of that line.

Vergara and Kim (2009) proposed a new method for buffer allocation in serial production lines. Their method can be implemented in a spreadsheet and the results were better than the compared genetic algorithm results. The production line was run using simulation and the results were compared with network for placing buffers. Metrics were developed in order to identify the buffer spaces which require more space based on the blocking, starving instances before each machine. Eight case studies with different scenarios such as balanced, unbalanced, reliable, automated, multiple bottlenecks were analyzed. The heuristic for buffer placement is shown in Figure 2.5.

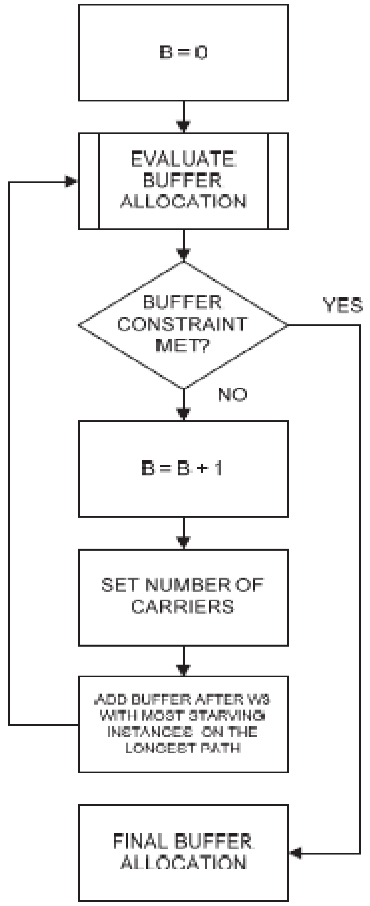

Figure 2-5 Flowchart for steps in buffer allocation (Vergara and Kim, 2009)

Shi and Gershwin (2009) developed an algorithm based on a non linear approach considering both buffer space cost and inventory cost with a minimum required throughput constraint. Battini, Persona, and Regattieri (2009) proposed a new paradigm "*buffer design for availability*". They handled the micro breakdowns in the production system by allocating buffers and thereby increasing the reliability of the system. They developed new guidelines for designing and allocating buffers in a system. The various reasons for the need of buffers are shown in Figure 2.6.

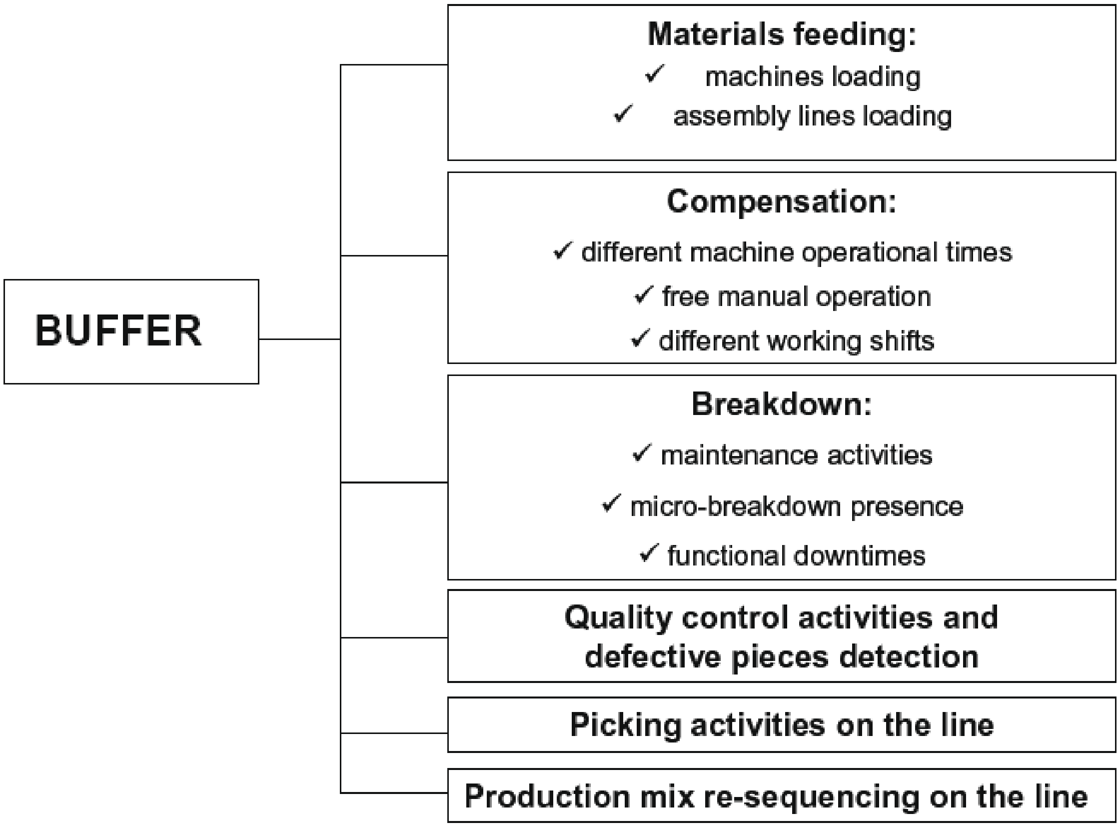

Figure 2-6 Functions of buffer (Battini et al., 2009)

Nahas, Ait-Kadi, and Nourelfath (2009) formulated a design on buffer and parallel machine allocation in unreliable production lines to increase its throughput. Their objective was to increase the throughput with a fixed total cost allotted for machines and buffers available in the market. The buffers are selected based on cost and size while the machines based on failure rate, repair rate, cost and size.

2.5 Buffer Allocation in Multi Product Production Systems

Nieuwenhuyse, Vandaele, Rajaram, and Karmarkar (2007) developed a queuing model for determining the buffer size for the required service level in a multi product multi reactor batch processing environment. The source of variability is from demand, setup and processing time of the products and they are considered to be stochastic. The system needed a minimum number of products for each process in a batch. So this buffer allocation method based on queuing model considered the tradeoff for a better solution.

Ye and Han (2006) developed a method for optimal size for stock buffer in front of a bottleneck station in a multi product production environment. The case studies tested in this paper had a single bottleneck station and the product routes were similar to the single bottleneck stations. So, actual multi product systems with alternate routes were not considered in this system.

2.6 Conclusion

From the above literature, it is clear that there is not enough work done for allocating buffer in multi- product and complex production systems. Previous work related to multi product systems doesn‟t have case studies involving actual multi-product systems. The mathematical models given by the researcher needs too many input data from the user. So there is a need for developing a real time allocation method where the user can allocate buffer based on the current situation of the system.

CHAPTER 3

METHODOLOGY

In this chapter, bottlenecks in a complex production system are identified using metrics developed based on various factors. Bottlenecks can be caused due to imbalanced production lines and variability in the system. Bottlenecks can be mitigated by using buffers and increasing the capacity of machines. Adding buffers are typically cheaper than increasing the capacity of bottleneck machines. However, in some situations machine capacity addition are needed to mitigate bottlenecks. The type of approach that is adopted depends on multiple factors that exist in the system. This chapter explores the type of bottlenecks, the metrics that have to be measured to determine the type of mitigation effort, and validates the method by using statistical analysis and simulation. These methodologies can be applied to the existing production systems for improving throughput. Further, the methodology for economic optimization of the system is derived based on sensitivity analysis.

3.1 Notations

The following notations are used in the methodology for finding buffer size and for adding capacity.

$\begin{pmatrix} P_{11} & ..... & P_{1k} \\ : & ..... & : \\ P_{j1} & ..... & P_{jk} \end{pmatrix}$	Process machine matrix, where P_{jk} is the process „j‟ at machine „k‟
$M_{V=1,,2,.....k}$	Machine 'M_V' , $V \in (1,2,......k)$
$B_{L=1,2,...N}$	Buffer locations based on variability (high to low)
$q_1, q_2 ......, q_N$	Size of buffer at each location „L‟

$S_{1,2,\ldots..i}$	Product 'S' , $S \in (1,2,\ldots\ldots i)$
T	Initial Throughput of the system
$M_{u1,\, u2,..uk}$	Machine arrangement utilization based (M_{u1} having highest utilization)
BT_V	Blocking time of machine „V‟
WT_V	Waiting time of machine „V‟
T_R	Required throughput
T_E	Economic throughput
T_N	Throughput after allocation
C_b	Cost per buffer space
P	Profit per product produced
T_{dec}	Throughput decrement after buffer removal
U_{avg}	Average utilization
X_V	= 0 if machine waits for a part for processing = 1 if machine doesn‟t need a part for processing
PT_{SV}	Processing time of product „S‟ at machine „V‟
Q_S	Quantity of product „S‟
a_{SjV}	Minimum processing time of product „S‟for process „j‟ at machine „V‟
b_{SjV}	Mode of the processing time of product „S‟for process „j‟ at machine „V‟
c_{SjV}	Maximum processing time of product „S‟for process „j‟ at machine „V‟
ρ_{SV}	Proportion of product „S‟at machine „V‟
σ_{SV}	Variance of processing time of product „S‟at machine „V‟

3.2 Bottleneck Identification

The throughput (T) of a production system can be improved either by adding capacity or by process improvement. There are several areas for process improvement such as scheduling, sequencing, reduction of cycle time, and eliminating non-value added time. This research focuses on reducing non-value added time by identifying and mitigating bottlenecks. Most of the time during non-value added state, the products wait before a bottleneck machine. Previous researchers have different definitions for a bottleneck machine.

According to Roser, Nakano, and Tanaka (2002) a machine which slows down the whole system is a bottleneck machine. There are different bottleneck identification methods as follows. According to Law and Kelton (2000), the machine with highest utilization is a bottleneck and Lawrence and Buss (1995) proposed that machine with the longest queue in its buffer is a bottleneck and Kuo, Lim, and Meerkov (1996) proposed the bottleneck based on sensitivity ratio. Tamilselvan, Krishnan and Cheraghi (2010), proposed the active duration method for bottleneck detection. These methods are tested in this chapter to determine their ability to predict bottlenecks in complex production systems. Although there are several researches related to bottlenecks, none defined bottlenecks for a multi product system with variability.

3.2.1 Identification Metrics

The machines which have higher risk of being bottlenecks that needed buffers or additional capacity can be easily identified with the developed metrics based on the processing time, number of products, and other performance parameters such as utilization, value added ratio, and queue length before machines. Previous research used high utilization and longest queue length separately for identifying bottlenecks.

3.2.1.1 Existing Strategies

There are several strategies that currently exist for identifying bottlenecks in single product cases. These methods may not be sufficient by itself to detect bottlenecks in complex production systems.

Longest Queue Length in Complex Manufacturing: Consider the job shop system shown in Figure 3.1. In this, machines 1, 2, 3, 4, 5, 6, 8, 9, and 10 are process operations in which a single part is processed at a time. Stations 7 and 11 are assembly stations. The product routing is given in the Table 3.1 and the processing time at each station is provided in Table 3.2. Each process has a different processing time as shown in Table 3.2. Parts A, B, and C undergo assembly operations in station 7 and further work on the assembled subcomponent which has parts A, B, and C is performed in Station 11. A simulation model was developed for this case study and after a run time of 1000 minutes, the results were analyzed.

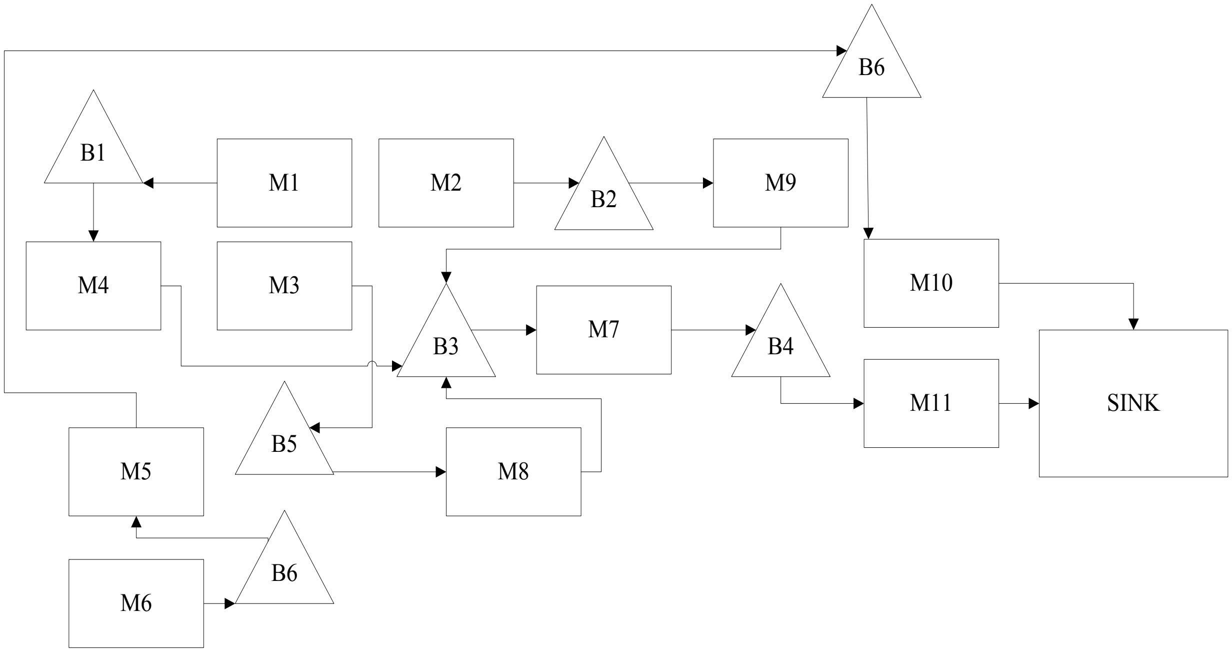

Figure 3-1 Layout of complex production system for Case Study 1

Table 3-1 Product Routing

Parts	**Routing**
A	1-4-7-11
B	2-9-7-11
C	3-8-7-11
E	6-5-10

Table 3-2 Processing Time for Case Study 1

Machines	1	2	3	4	5	6	7	8	9	10	11
Part	A	B	C	A	E	E	ABC	C	B	E	ABC
Min	3	4	5	5	3	5	4	5	5	4	4
Mode	7	10	7	10	5	6	5	7	13	6	5
Max	10	13	10	15	9	7	6	10	17	7	6

Table 3-3 Performance Analysis of Case Study 1 Results

Machines	1	2	3	4	5	6	7	8	9	10	11
Utilization	99.4	98.4	100	99.2	93.8	100	37.7	97.3	98.9	92.2	37.5
Average Queue Length	2	3	5	20.6	0.29	4	40.5	0.73	15.1	.17	0

The results of the simulation are shown in Table 3.3. In this system, Machine 7 has the highest average queue length of 40.5 parts. Since machine 7 needs Part A, B, and C for assembly, the unavailability of one type of part would make the machine idle which in turn decreases the utilization of the station and builds a large queue. Since, this queue length consists of parts A, B, and C and it is misleading to conclude that the queue is long. The next highest queue length (20.6) in the system is at Machine 4. However, machine 4 is not a bottleneck. Since the production rate of Machine 4 is high compared to other machines and there are enough products flowing through Machine 4 to Machine 7 for assembly, it can be concluded that Machine 4 is not a bottleneck. Thus, the longest queue length may not always be a bottleneck. After analysis of the simulation results, it is concluded that Machine 9 is the dominant bottleneck in the system. This conclusion is based on the bottleneck definition proposed by Roser, Nakano, and Tanaka (2002) which states that a machine which slows down the whole system is a bottleneck machine. Machine 9 has a queue length of 15.1, which is not the longest in the system. This machine starves the Machine 7 and 11 from assembling. Thus the longest queue length criteria may not be sufficient by itself to track bottleneck machines in complex production systems. However, it is possible that the longest queue length could be used with other metrics to determine the bottleneck.

Highest Utilization in Complex Manufacturing: In the case study 1, Machines 1, 3, and 6 have the highest utilization of 99.4%, 100%, and 100% respectively (Table 3.3). But all three machines meet the demand and have a high production rate. Machine 9 has a utilization of 98.9% which is less than that for machines 1, 3, and 6. Machine 9 is the bottleneck machine. From this, it can be seen that the highest utilized machine may not always be the bottleneck in a complex production system.

To further invalidate the highest utilization method, another case study (Figure 3.2) for a complex system was developed. The development of a case study with no buffers was done to eliminate the effect of buffer on utilization. The product sequence for this case study is that same as the one in the first study. The processing times for this case study are shown in Table 3.4. The results of the case study after simulation are provided in Table 3.5. The machines with the highest utilization are Machine 6, 5, and 10 at 96.6%, 95.7% and 95.6% respectively. But these machines are not bottleneck machines and do not slow down the system. The bottleneck machines are 2 and 9 which fail to supply parts to the assembly station. The parts that are processed in 2 and 9 have processing time, higher variability, and the inter-arrival rate for parts is high. Thus, in a complex production system, the highest utilization machine in a system may not be the bottleneck machine.

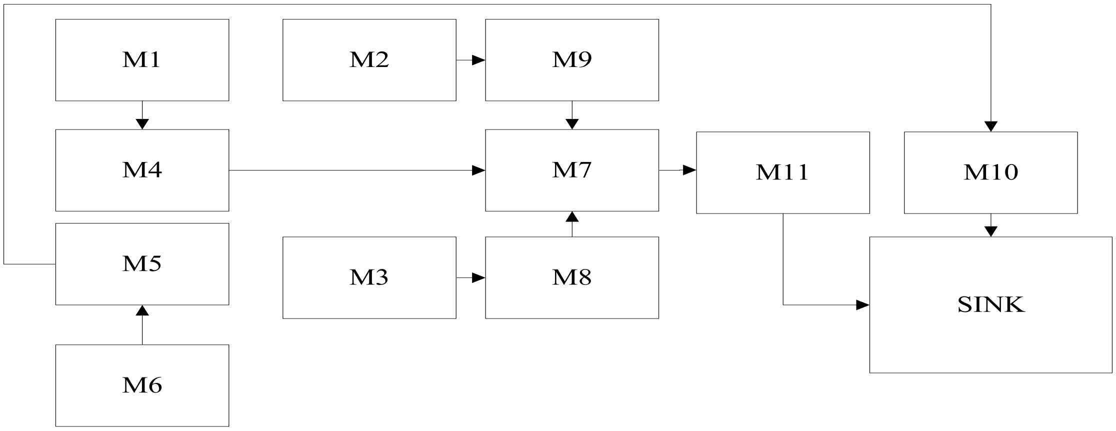

Figure 3-2 Layout of complex production system for Case Study 2

Table 3-4 Processing Time for Case Study 2

Machines	1	2	3	4	5	6	7	8	9	10	11
Part	A	B	C	A	E	E	ABC	C	B	E	ABC
Min	3	9	5	5	5	5	4	5	10	5	4
Mode	7	14	7	10	5	5	5	7	20	5	5
Max	10	18	10	15	6	6	6	10	21	6	6

Table 3-5 Performance Analysis of Case Study 2 Results

Machines	1	2	3	4	5	6	7	8	9	10	11
Utilization	47.9	91.6	49.9	69.7	95.7	96.6	33.9	50.7	90.2	95.6	33.9

3.2.1.2 Modified Bottleneck Identification Metrics

Identification of a bottleneck in a multi-product system in the presence of variability is a complex problem. As the bottlenecks do not follow any pattern, it is difficult to develop a single generic method/measure to find the bottlenecks. The variability in a machine either causes blocking or starving in upstream and downstream machine. Starving occurs if a machine is waiting for a part from an upstream machine and blocking occurs if a downstream machine blocks the flow of products in the system. So the bottleneck stations have to be improved in order to avoid blocking or starving. In complex production systems, blocking and starving occurs instantaneously, which is complex to predict. New metrics are needed to identify the risk for a machine to become a bottleneck. In addition, the utilization metric and the queue length metric have to be modified for the complex production system. These modified metrics and the new metrics are defined below. The highest utilization and longest queue length is combined for identifying the bottlenecks accurately. Each machine"s utilization and queue length is compared with the average queue length and average utilization of the system to identify the bottlenecks.

Modified Utilization Metric (U_R): In this metric, the bottleneck machine is found based on the utilization. Once the U_{avg} is found, it is compared with the utilization of each machine. If the utilization ratio of a machine is larger than the average utilization, it may be a bottleneck.

$$U_{avg} = \frac{\left(U_1 + U_2 + \ldots\ldots U_k\right)}{k} \tag{3.1}$$

$$U_V = \frac{\text{Total active time of machine 'V'}}{\text{Total time of the system}} \tag{3.2}$$

$$U_R = \frac{U_V}{U_{avg}} \tag{3.3}$$

Modified Queue Length Metric (W_V): In this metric, the machine with the longest queue length may be a bottleneck. Similar to the utilization metric, W_{avg} is found and it is compared with the „W" of each machine.

W_{avg} – Average WIP

$$W_{avg} = \frac{W_1 + W_2 + \ldots\ldots\ldots + W_L}{k-1} \tag{3.4}$$

$$W_V = \left[\frac{\text{WIP before machine 'V'}}{\text{Total WIP in the system}}\right] * X_V \tag{3.5}$$

3.2.1.3 Additional Bottleneck Identification Metrics

Variance Metric (γ_V): The processing time of machines in complex production system exhibits two types of variability such as variability caused by the process itself and the variability caused by the difference in processing time between products. In order to capture the total variance of the machine the equations are used as shown below. The variance formula is selected for the triangular distribution which is used in our research.

$$\text{Total Variance of a machine 'V'} = \sqrt{\sum_{S=1}^{i} (\rho_{SV} * \sigma_{SV}^2)} \tag{3.6}$$

$$\text{Variance of a processing time for product 'i' in machine 'k' } (\sigma_{iV}^2) = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18} \tag{3.7}$$

Balancing Metric (ε_V): The degree of imbalance between machines due to the varying processing times is determined in this metric. The ratio of products processed in each machine is considered along with the processing time to find the degree of imbalance between machines.

The buffer level can be limited if there is a high imbalance between machines as the high buffer decreases the total system performance.

$$\varepsilon_V = \frac{\sum_{S=1}^{i} (PT_S Q_S)}{\sum_{S=1}^{i} Q_S}, \quad V=1,2......k \tag{3.8}$$

Process Time Metric (β_{SV}): In a multi product production system, product „S‟ might cause machine „V‟ to be a bottleneck, if it needs more processing time than other products. This metric would determine the bottleneck process „j‟which processes product „S‟ at machine „V‟.

$$\beta_{SV} = \frac{\text{Processing time of product 'S' in machine 'V'}}{\text{Total processing time of a machine}} \tag{3.9}$$

Previous methods found the bottlenecks using the results from simulation and the heuristics for control strategies were also developed based on simulation approach.

Apart from finding a bottleneck, the type of bottleneck should be determined based on the type of control strategies allotted. There are two types of bottlenecks based on the type of variability. They are,

- Variance bottlenecks caused by processing time variability within the machine
- Mean bottlenecks caused by difference in mean processing time between machines

The bottlenecks in a complex production system are identified using metrics developed based on various performance parameters. Since the machine with the highest utilization in the system has a tendency to become a bottleneck, the metrics are developed on this parameter along with supporting factors to find an accurate solution. Once the bottlenecks are identified, the type of bottleneck is defined based on other metrics and the respective control strategies are implemented. If the bottleneck is variance type then buffers are allotted initially and if it is a mean bottleneck caused by an imbalance between machines, additional capacity is used before allotting buffers. The need for optimal buffers and additional capacity is detailed below.

This research also focuses on machines with processing time variability and so the bottlenecking effect may last only for a short period. This can be handled effectively by placing buffers between machines. Infinite buffer always have shown improved throughput which in turn decreases the total system performance. As there is an increase in buffer space, there is an increase in holding cost and the work in process (WIP) inventory. The main goal for any industry is to increase profit. So, the process improvement or buffer allocation has to be economically viable rather than concentrating on production rate and throughput.

3.3 System Evaluation Metrics

Tamilselvan, Krishnan, and Cheraghi (2010) proposed four metrics for identifying bottlenecks in single product production systems. These metrics could help in identifying the impact and the type of bottlenecks. The metrics are a) Bottleneck time ratio, b) Bottleneck ratio, c) Bottleneck shifting frequency and d) Bottleneck severity ratio. It is impossible to determine the shifting frequency in a complex production system manually, as the shifting occurs instantaneously at each and every moment in the presence of variability. In this research, the metrics applicable to complex production systems are bottleneck time ratio, bottleneck ratio, and value added ratio.

3.3.1 Bottleneck Time Ratio

The bottleneck time ratio determines the total bottleneck time in the system. This metric would determine the efficiency of the system based on the bottleneck time.

$$\alpha = \frac{\text{Bottleneck Time}}{\text{Total Run Time}} \tag{3.10}$$

3.3.2 Bottleneck Ratio

Bottleneck ratio determines number of bottleneck machines in the system.

$$\tau = \frac{\text{Number of Bottleneck Machines}}{\text{Total number of Machines}} \tag{3.11}$$

3.3.3 Value Added Ratio

Apart from these proposed metrics, Karthikeyan (2010) used value added ratio metric.

$$\text{VAR}=\frac{\text{Active state time}}{\text{Inactive state time}} \tag{3.12}$$

In addition to the above three metrics, throughput and machine utilization will also be used to verify the improved performance. These metrics are developed based on single product production systems. The above metrics can also be used in complex production systems for measuring the efficiency of the system. In addition to the above defined metrics, additional metrics are needed to comprehend the bottleneck problem. These metrics are explained in the following section of this chapter.

3.4 Bottleneck Mitigation Strategies

Once the bottlenecks are found using the above metrics, they need to be classified based the type of variability. In this research, two types of bottlenecks are analyzed in complex production systems. They are variance bottlenecks and mean bottlenecks. Variance bottlenecks are caused by the variability in processing time and the mean bottlenecks are caused by imbalance in processing time between machines. In order to mitigate the bottlenecks effectively, the type of bottleneck has to be determined before adopting control strategies.

There are methods to mitigate bottleneck such as additional capacity and buffer allocation. Adding capacity is usually expensive and needs more storage space. But if the bottleneck is caused by imbalance between mean processing times in machines, there is no use in allocating buffer in between machines. So the type of bottleneck can be identified from the imbalance metric. Buffer allocation is a traditional and economic option which also has an extensive research since 1960. There are many generalizations in buffer allocation in the previous research. Some important characteristics can be adapted to mitigate complex production systems. The two methods for mitigating bottlenecks are explained in detail as follows.

In this section, simulation based heuristics are developed for handling bottlenecks and thereby improving the performance of the system. The objective of this section is three fold as follows.

- Buffer allocation based on the qualitative characteristics
- Capacity addition based on highest utilization method
- Economic analysis based on sensitivity analysis

3.4.1 Optimal Buffer Allocation

3.4.1.1 Qualitative Characteristics of Buffers

Previous researchers studied on qualitative properties of a production system for buffer allocation and made generalizations based on the number of machines in the system and buffer size constraints. In Figure 3.2, a typical plot of throughput vs. buffer size is provided. The curve can be divided into two major regions: continuous and discontinuous. According to Gershwin and Schor (2000), the qualitative properties in buffer allocation are continuity, monotonicity, and concavity, and these apply to the continuous region.

Apart from these properties there are many generalizations for different systems. Most of the generalizations can only be applied to specific systems and it could not be considered for complex multi product systems

Transition Point

In every manufacturing system, there exists a transition point that divides the continuous region from the discontinuous region. This point is at which there is no increase in throughput for buffer increase $\left(\frac{dT}{dB}=0\right)$

Continuity Property

Every increase in buffer size within the continuous region has an impact on system performance (Gershwin and Schor, 2000). In the presence of machines with high variability in processing time, buffers dampen the variability. Thus, with each change in the buffer, the system throughput changes within the continuous region.

Monotonicity

Any increase in buffer size increases the production rate, as stated by Gershwin and Schor (2000). i.e., the curve plot of $\left(\frac{dT}{dB}=0\right)$ is continuous and the first derivative exists within the continuous region and it maintains the same sign until the transition point.

Concavity

The percentage increase in throughput decreases, when there is a continuous increase in buffer size (Gershwin and Schor, 2000). The rate of change in the performance will either be continuously decreasing or increasing with reference to any point in the continuous region. The

$\left(\frac{d^2T}{dB^2}=0\right)$ will experience a continuous decrease for each unit increase in buffer. At the

transition point, the $\left(\frac{d^2T}{dB^2}=0\right)$.

This property holds good in most of the cases and even some cases exhibit very little improvement for an increase in buffer size. The necessity for placing or removing the buffer can be justified with this property. This property is shown in the Figure 3.4.

Discontinuity

An increase in buffer size would have no effect on the system throughput in the discontinuous region. In such cases, an addition of buffer decreases the total system performance because of factors such as increase in holding cost and high non value added time for products.

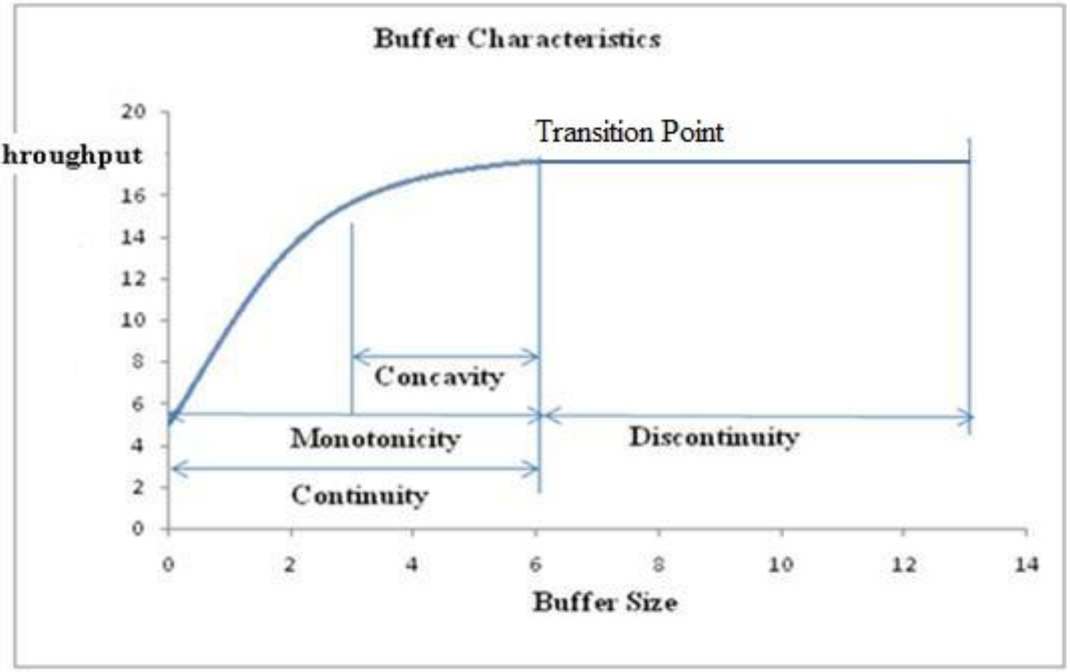

Figure 3-3 Qualitative Properties of Buffer Allocation

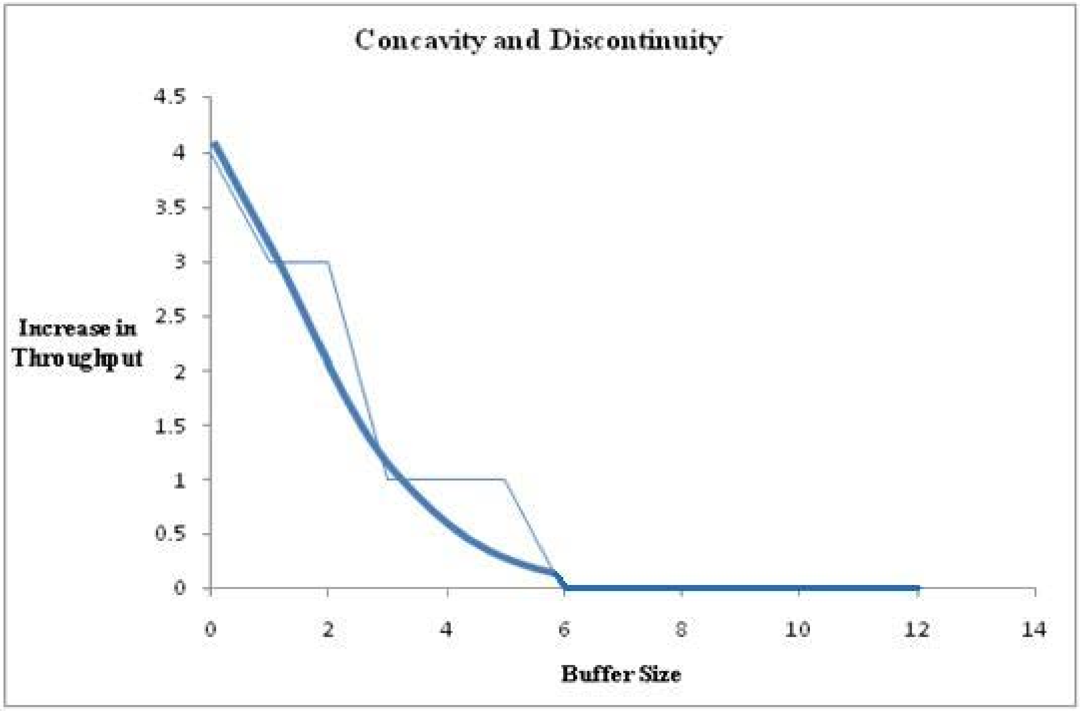

Figure 3-4 Concavity and Discontinuity

Most of the mixed model assembly lines and automated production systems such as air conditioning units, engine manufacturing, and printer assembly have layouts similar to the case studies studied. In these systems, even though the products processed have same routing, and precedence constraints there is a difference in processing time for products with different specifications. Since the mathematical calculation is very tedious and cannot be applied to different scenarios, a heuristic method is developed. The layouts with sequential flow and jumbled flow can be solved using this methodology.

3.4.1.2 Methodology for Buffer Addition Analysis

The optimal buffer for a multi product system is determined using the heuristic based on longest queue length method and validated using simulation. The first step of the heuristic is to determine the maximum possible throughput in the system by allocating infinite buffer to the highest variability station (Variance Metric). This maximum throughput acts as an upper bound and the optimal buffer should have a throughput equal to the upper bound.

The optimal buffer size is attained by reducing the buffers at each buffer location. The algorithm is terminated once the final buffer location is optimized. The flow chart for the buffer allocation is shown in Figure 3.5. If there is a high imbalance because of a dominant bottleneck machine, even the infinite buffer before machines would result in a reduced throughput. If the reduced throughput is less than or equal to the required throughput, there is a need to find a solution to handle the dominant bottleneck. The machine allocation methodology shown in Figure 3.7 is used based on the blocking time and waiting time of the upstream and downstream machines for the highest utilization machine.

3.4.1.3 Algorithm

```
Step1:  Start
Step2:  Start the line by assigning buffers $B_{L=1,2,....N}=\infty$ before all the machines.
Step3:  The throughput obtained for the infinite buffer is $T_{max}$.
Step4:  Go to $B_{L=1}$. (Machine with the highest variability)
Step5:  Keep reducing the buffer size at $B_L$ until $T_{max}$ reduces.
Step6:  Increase the buffer size by one and then move to $B_{L+1}$ location (machine with next
        high variability).
Step7:  Repeat from step 5 to 6 until L>=N.
Step8:  Stop.
```

3.4.1.4 Flow Chart

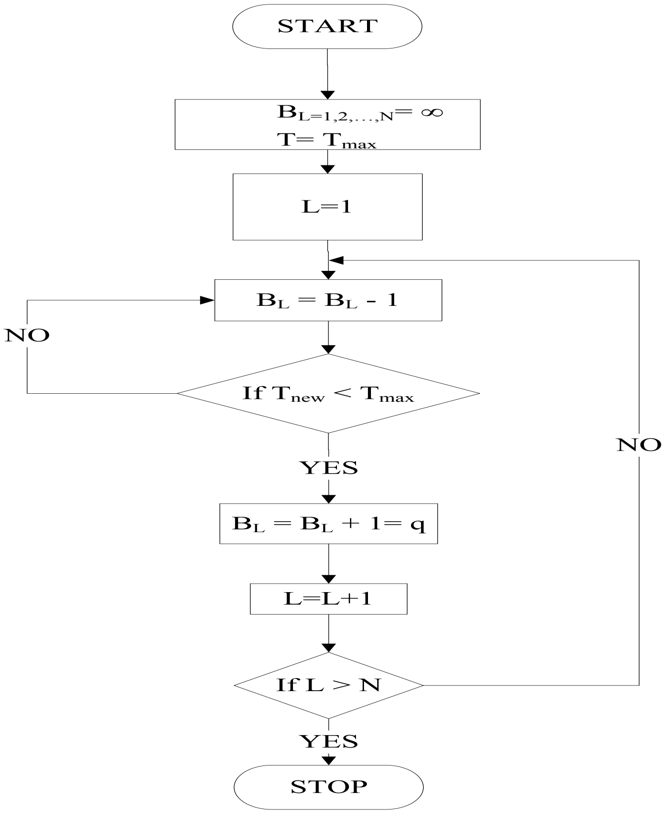

Figure 3-5 Flow chart for optimal buffer in single line system

3.4.1.5 Example Case Study

In this case study, a system shown in Figure 3.6 with 3 products and processing time shown in Table 3.4 is selected to explain the buffer methodology. This case can be extended to larger case studies but the selection of buffer locations is in ascending order based on the variance metric value.

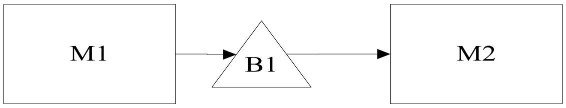

Figure 3-6 Example case study for optimal buffer method

Table 3-6 Processing time for 2 Machine and 3 Product case study

	Machine 1			Machine 2		
	Min	**Mode**	**Max**	**Min**	**Mode**	**Max**
Part 1	20	25	40	15	20	35
Part 2	10	15	25	20	45	45
Part 3	25	30	50	10	15	25

Based on the method proposed, the buffer sizes were reduced from infinite to the maximum utilized and compared with the throughput rate. If there is a throughput reduction because of the buffer reduction, the decrement in buffer size is terminated. The optimal buffer is determined as shown in Table 3.5.

Table 3-7 Optimal buffer methodology

Buffer 1	Average throughput
Infinite	936.2
6	936.2
5	936.2
4	936.2
3	933.2
2	911.1

3.4.2 Capacity Addition Method

Even though the buffer allocation increases throughput, if the bottleneck machine is severe and if the system is not meeting the required throughput, additional capacities are needed to meet the required demand.

3.4.2.1 Methodology for Additional Capacity Analysis

Based on the utilization, blocking time of upstream machine and waiting time of downstream machine, the bottleneck machines are identified for adding resources. The flow chart for this method is shown in Figure 3.7 and the method is explained using a sample case study.

3.4.2.2 Algorithm

Step1: Start

Step2: Arrange the machines in descending order based on utilization Now ($M_{V=1,2,\ldots..K}$) is ($M_{u=1,2\ldots K}$)

Step3: Starting from the highest utilization machine, consider the corresponding „M_V"

Step4: If $BT_{k-1} \geq BT_k$ and $WT_{k+1} \leq WT_K$ then add a capacity to the workstation.

Step5: Verify the new throughput (T_{new}) with the required throughput (T_{req})

Step6: Repeat step 2.

Step7: Else go to the next machine based on utilization and repeat steps 2 and 3 until u>k.

Step8: Stop.

3.4.2.3 `Flowchart

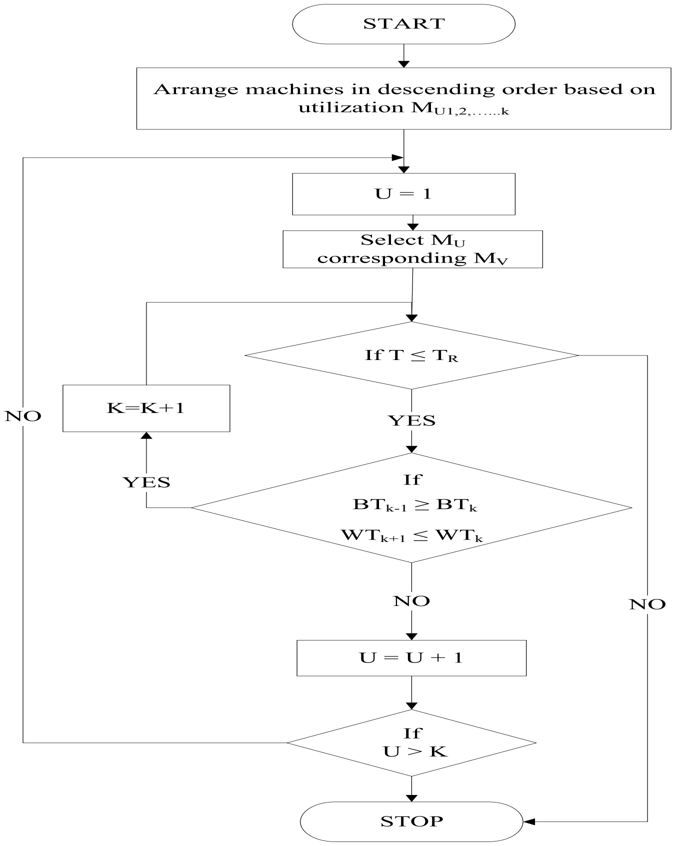

Figure 3-7 Flow chart for adding machines to meet the required throughput

3.4.2.4 Example Case Study

In this case study, 3 machines and 3 products system as shown in Figure 3.8 is considered for verifying the methodology developed for capacity addition. The processing time table is shown in Figure 3.6. The simulation is run for 25000 minutes and the performance results before and after capacity addition is shown in the Table 3.7 and Table 3.8.

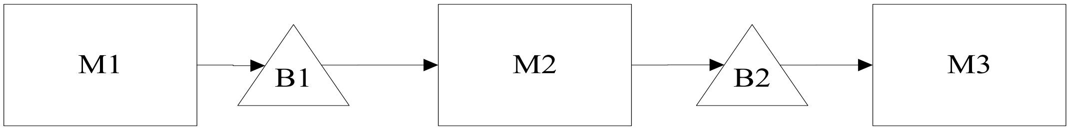

Figure 3-8 Example case study for additional capacity before allocation

Table 3-8 Processing time for 3 Machine/3 Product case study

	Machine 1			Machine 2			Machine 3		
	Min	**Mode**	**Max**	**Min**	**Mode**	**Max**	**Min**	**Mode**	**Max**
Part 1	20	25	30	35	45	70	10	15	20
Part 2	10	15	20	30	45	45	12	15	18
Part 3	25	25	30	30	45	55	10	12	15

Before Allocation

Once the simulation is run for 25000 minutes, the machines are arranged in descending order based on their utilization. In this case study Machine 2 has the highest utilization and as per the method the blocking time of Machine 1 and Waiting time of Machine 3 is analyzed. Since both these values large when compared to Machine 2, it is considered a bottleneck.

Table 3-9 Performance parameters before capacity addition

	Utilization	Blocked Time	Waiting Time
Machine 1	50.50%	12355.22 min	18.33 min
Machine 2	99.89%	0 min	43.04 min
Machine 3	31.75%	0 min	17063.51 min

After Allocation

Since Machine 2 is bottleneck, an additional capacity is allocated to it and the simulation is run for analyzing the improvement. The blocking and waiting time of upstream and downstream machines are considerably and the throughput is improved from 855 to 1115.2 parts

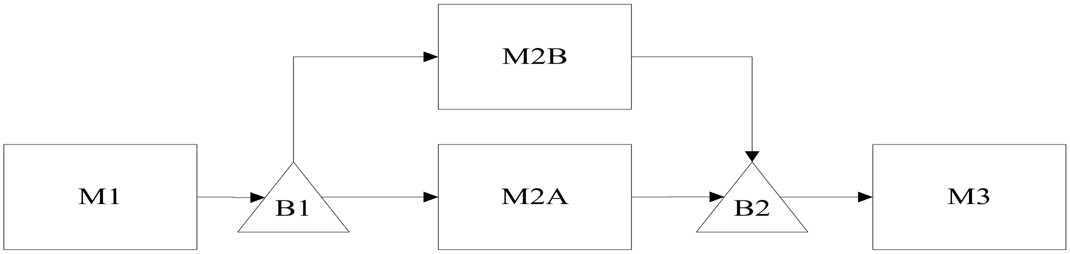

Figure 3-9 Example case study for additional capacity after allocation

Table 3-10 Performance parameters after capacity addition

	Utilization	Blocked Time	Waiting Time
Machine 1	99.67%	61.94 min	18.39 min
Machine 2A	99.29%	0 min	175.8 min
Machine 2B	99.21%	0 min	195 .83 min
Machine 3	63.05%	0 min	9235.49 min

3.4.3 Economic Analysis

In the previous section, the method for allocating optimal buffer has been discussed in detail. Now the buffer size should be minimized based on the economic feasibility. Though the buffers allotted gave the maximum possible throughput, it might not be economically feasible because of the concavity property. Now an economically feasible buffer can be derived from the method given below.

3.4.3.1 Methodology

In this section, a methodology for economic buffer is proposed for a multi product system. Even though the buffer found in previous sections were optimal, it might not be economical as some buffers would not have a significant impact on the throughput. In this method, the sensitivity of each buffer is analyzed by reducing the buffer further from the optimal throughput point. If there is no significant increase and if the holding cost of buffer is larger than the profit of throughput, the buffer is considered a loss and it is removed. The flowchart for the method is shown in Figure 3.10.

3.4.3.2 Algorithm

Step1: Start

Step2: The initial buffer size from optimal buffer method $B_{L=1,2,...N} = q_1,q_2,...,q_N$ and the respective throughput $T = T_R$.

Step3: Starting from the B_1 reduce buffer quantity q_1 by one and the new throughput (T_N) is noted.

Step4: Note the difference between T_N and the initial throughput (T_R). ($T_{dec} = T_R - T_N$).

Step5: If the profit from T_{dec} is greater than the buffer space reduced, add one buffer to B_1 and go to next machine.

Step6: Move to next buffer location (L= L + 1)

Step7: Repeat the steps until final buffer location

Step8: Stop

3.4.3.3 Flow Chart

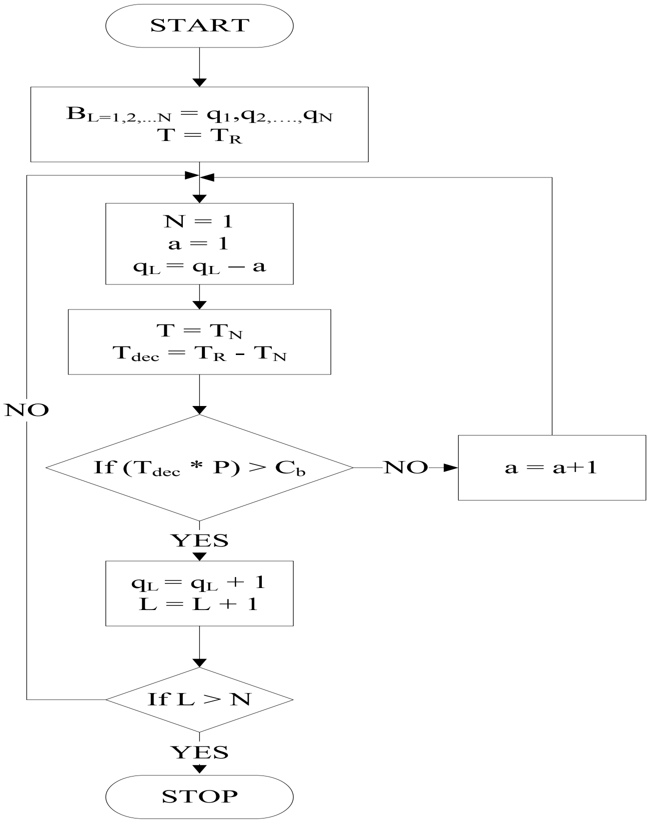

Figure 3-10 Flow chart for determining economical buffer

3.4.3.4 Example Case Study

The economic buffer allocation is explained using a 4 machine and 3 part case study as shown in Figure 3.11. Assuming the cost per buffer space as \$10 and profit per product as \$5 the economic feasibility is calculated and results are shown in Table 3.10.

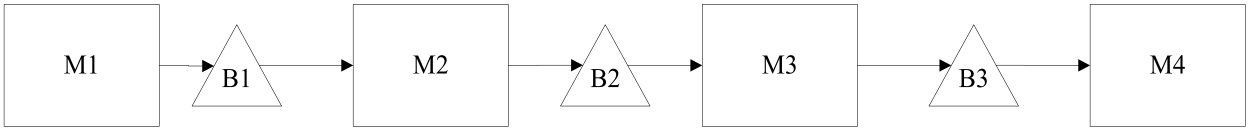

Figure 3-11 Example case study for economic buffer allocation

Table 3-11 Economic buffer analysis

	Buffer	**Throughput**
	(∞, ∞,∞)	248.6
Optimal	(3,2,1)	248.6
	(1,2,1)	248.2
	(0,2,1)	243
	(1,1,1)	247.4
	(1,0,1)	241.1
Economical	(1,1,0)	246.6

In this case, the maximum throughput is 248.6 parts with optimal buffer size (3, 2, and 1). The profit of the system for this allocation is \$1,183 after subtracting the expenses for buffer space. The buffer is reduced further based on the cost factor and checked for economic feasibility. Buffer size (1, 1 and 0) selected since it is economical and gives a profit of \$1,213 which is larger than the previous allocation.

CHAPTER 4

CASE STUDIES

4.1 Simulation Assumptions

All case studies are simulated using discrete event simulation software Delmia Quest V5. The processes are simulated for 25000 minutes and replicated 10 times in order to capture the effect of variability as accurately as possible. The inputs of the products are kept constant at all times to find the improvement in throughput irrespective of the difference in processing times of different products. The various assumptions are as follows:

- Processing time for each product in each machine follows a triangular distribution
- FIFO system is used for processing orders
- Machines are available at all time without any failures
- All products have their own product route with precedence constraint
- Buffers are placed between two machines
- The last machine is never blocked
- Material handling time is zero and distance are also zero
- Cost/buffer = \$10, Cost/capacity = \$2500

4.2 Scenarios Based on Bottleneck Metrics

The proposed metrics (Chapter 3) can be used to identify the actual bottleneck and the source which causes it. Based on the values of each metric, the behavior of each machine is identified and the control strategies needed for improving the performance is assessed. This is further validated by simulation based heuristics and economic analyses. The possible combinations (scenarios) that can occur in the system based on the metrics are shown in Table 4.1.

Table 4-1 Possible Combinations of Metrics by Design of Experiments

Utilization Metric	Queue Length Metric	Imbalance Metric	Variance Metric
Low	Low	Low	Low
High	Low	Low	Low
Low	High	Low	Low
High	High	Low	Low
Low	Low	High	Low
High	Low	High	Low
Low	High	High	Low
High	High	High	Low
Low	Low	Low	High
High	Low	Low	High
Low	High	Low	High
High	High	Low	High
Low	Low	High	High
High	Low	High	High
Low	High	High	High
High	High	High	High

Some of the combinations which are not feasible and impossible to occur in complex production systems are eliminated in steps.

- Since, imbalance metric and queue length metric are directly proportional; all the possibilities where they are indirectly proportional are removed (8 Possibilities)
- If there is long queue before machine „V", then the utilization of machine „V" cannot be low (2 possibility)
- Apart from these systems, there are six scenarios where there are chances for machines to become bottlenecks. These are shown in Table 4.2

Table 4-2 Realistic Scenarios

Scenarios	Utilization Metric	Queue Length Metric	Imbalance Metric	Variance Metric	System Type
1	High	High	High	Low	Unbalanced
2	High	Low	Low	High	Balanced line with variability
3	High	High	High	High	Worst case scenario
4	Low	Low	Low	Low	Under utilized Machine
5	Low	Low	Low	High	Under utilized with variability
6	High	Low	Low	Low	Balanced line (Best System)

These scenarios are tested considering various case study settings.

4.3 Case Study 1

In this case study, the systems which are similar to single product production systems are considered. Most of the mixed model assembly lines have these kinds of systems. In these systems, the processes involved in processing the different products are the same but the processing time differs from each product as the specification of the products differ for each product. Some of the examples for these kinds of systems are printer assembly, automobile manufacturing, computer assembly etc.

Figure 4-1 Block diagram for case study 1 with 5 machines and 4 buffers

4.3.1 Product Routing

In this type of system, all the products have the same product routes and each product has a different processing time as shown in Table 4.3. In this system, five machines process all the products. Most of the mixed model assembly lines work similar to this system.

Product 1: M1-M2-M3-M4-M5

Product 2: M1-M2-M3-M4-M5

Product 3: M1-M2-M3-M4-M5

Table 4-3 Processing Time Table for Case Study 1

	Machine 1			Machine 2			Machine 3			Machine 4			Machine 5		
	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max
Part 1	80	120	150	60	140	150	100	125	150	100	135	150	100	120	150
Part 2	100	115	150	100	140	150	70	125	170	60	135	150	100	120	150
Part 3	100	135	150	100	145	150	100	135	150	100	145	150	100	130	150

4.3.2 System Evaluation Metrics

Bottleneck Time Ratio: In this system, the bottleneck time is comparatively low when compared to the total run time. So the bottleneck time ratio is also low.

$$\text{BN Time Ratio} = \frac{1350.4}{25000} = 0.054$$

Bottleneck Ratio: In this system, four machines were bottlenecks at least in one instant.

$$\text{BN Ratio} = \frac{4}{5} = .8$$

Value Added Ratio: The value added activity is very high when compared to the non value added activity and so the ratio value is 16.51.

$$\text{VAR} = \frac{22299.11}{1350.4} = 16.51$$

4.3.3 Bottleneck Identification Metrics

Utilization Metric (U_V): The utilization of each machine is compared with the average utilization of the total system as shown in Table 4.4.

$$U_{avg} = \frac{89.8+93.8+85.4+88.8+88.1}{5} = 89.2$$

Table 4-4 Utilization Metric for Case Study 1

	U_1	U_2	U_3	U_4	U_5
U_K	89.8	93.8	85.4	88.8	88.1
Ratio	0.99	1.04	0.96	1.01	0.97

Since the utilization ratio of Machine 2 is large when compared to other machines, the chance for Machine 2 becoming bottleneck is high.

Queue Length Metric (W_V): The queue length of each machine is compared with the average queue length of the system as shown in Table 4.5

Average Queue Length = 3.45

Table 4-5 Queue Length Metric for Case Study 1

	W_1	W_2	W_3	W_4	W_5
Ratio	0.29	2.84	0.31	0.55	0.29

The queue length before Machine 2 (W_1) is larger when compared to other machines and since both the bottleneck identification metrics identify the same machine as a bottleneck, it can be concluded that Machine 2 as a bottleneck.

Balancing Metric (ε_V): The degree of imbalance is low in this system as shown in Table 4.6 and it can be solved by adding buffers instead of additional capacity.

Table 4-6 Degree of imbalance for Case Study 1

ε_1	ε_2	ε_3	ε_4	ε_5
0.98	1.01	1.00	1.00	1.00

Variance Metric: The variability in the processing time is high in Machine 3 and Machine 4 which is shown in Table 4.7 using the variance metric.

Table 4-7 Metrics based on variability for Case Study 1

γ_1	γ_2	γ_3	γ_4	γ_5
10.4	10.9	14.4	14.3	10.2

Processing Time Metric (β_{SV}): In this system, the processes involving Part 2 in Machine 2 and Machine 3 have high processing time when compared to other machines and the values are shown in Table 4.8. The transition between Part 2 and other parts at Machine 3 and 4 might need high buffer space when compared to other buffer locations.

Table 4-8 Processing Time for Case Study 1

	Machine 1	Machine 2	Machine 3	Machine 4	Machine 5
Part 1	$\beta_{11} = 0.84$	$\beta_{12} = 0.89$	$\beta_{13} = 0.83$	$\beta_{14} = 0.88$	$\beta_{15} = 0.84$
Part 2	$\beta_{21} = 0.88$	$\beta_{22} = 0.89$	$\beta_{23} = 3.34$	$\beta_{24} = 3.1$	$\beta_{25} = 0.84$
Part 3	$\beta_{31} = 0.88$	$\beta_{32} = 0.96$	$\beta_{33} = 0.88$	$\beta_{34} = 1.0$	$\beta_{35} = 0.84$

4.3.4 Control Strategies

Bottleneck machines and the type of bottlenecks are identified using metrics as shown in Table 4.9. Since the imbalance between machines is low and the low imbalance is caused by the variability of the system, buffer allocation is the efficient method to mitigate bottlenecks. Buffers are also the cheapest method when compared to the other possible methods as it has no maintenance and can be removed at any time unlike additional capacities.

Table 4-9 Comparison of Metrics

Metrics / Machines	Utilization Metric	Queue Length Metric	Balancing Metric	Variance Metric
Machine 1	0.99	0.29	0.98	10.4
Machine 2	1.04	2.84	1.01	10.9
Machine 3	0.96	0.31	1.00	14.4
Machine 4	1.01	0.55	1.00	14.3
Machine 5	0.97	0.29	0.99	10.2

The imbalance between the machines is low which is validated by queue length metric, utilization metric, and imbalance metric. Since the variability metric values are high, there is a need for allocating buffers in order to improve the performance of the system. Based on the method developed which is shown in Chapter 3, buffers are allocated to the system as shown in Table 4.10. The initial allocation of buffers (3, 1, 2, and 1) gives an optimal throughput and further the buffers are reduced based on the economic analysis. The buffer spaces (2, 1, 1, and 1) which are more economically feasible can be selected based on the management‟s decision. The performance improvement of the system is improved as shown in Table 4.11

Table 4-10 Optimal buffer for case study 1

	Buffer	Average throughput
	Infinite	186.6
	(10,2,3,1)	186.5
Optimal	(3,1,2,1)	186.5
Economical	(2,1,1,1)	186.1

Table 4-11 Performance Measures for Case Study 1

	Before	After
BN Time Ratio	0.054	0.014
BN Ratio	0.8	0.2
Value Added Ratio	16.51	24.1
Throughput	176	186.1

4.4 Case Study 2

The above model is validated using a simple 6 machines and 3 products case study as shown in Figure 4.2 In this system the minimum required throughput is 235 parts and the buffer allocation should be based on this condition. If the required throughput could not be reached, the methodology for adding another machine is used to satisfy the condition. The processing time for the system is shown in Table 4.12.

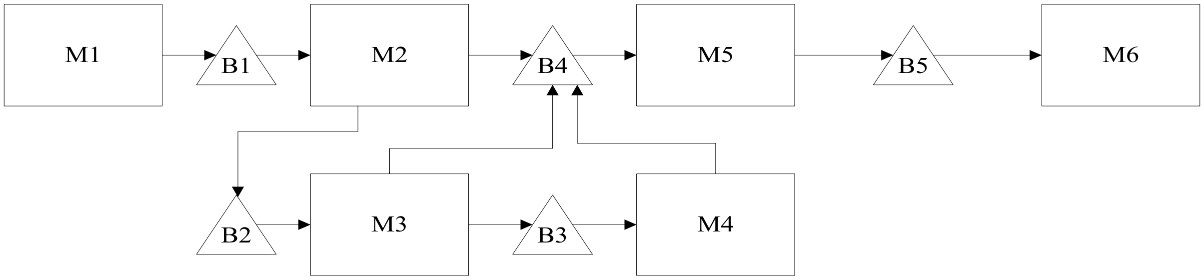

Figure 4-2 Block diagram for case study 2 with 6 machines

4.4.1 Product Routing

In this type of system, each product has a different route and processing time as shown below. In this system, four machines process all the products and the throughput depends on the machine with highest processing time.

Product 1: M1-M2-M3-M5-M6

Product 2: M1-M2-M5-M6

Product 3: M1-M2-M3-M4-M5-M6

Table 4-12 Processing Time Table for Case Study 2

	Machine 1			Machine 2			Machine 3		
	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max
Part 1	75	110	120	75	100	120	90	120	140
Part 2	75	115	120	75	95	120			
Part 3	75	100	120	75	100	120	90	130	140
	Machine 4			Machine 5			Machine 6		
	Min	Mode	Max	Min	Mode	Max	Min	Mode	Max
Part 1				90	120	120	75	90	100
Part 2				90	110	120	75	85	100
Part 3	120	140	150	90	115	120	75	90	100

4.4.2 System Evaluation Metrics

The efficiency of the system is measured using the three different metrics and it is again measured after implementing the control strategies for analyzing the performance improvement.

Bottleneck Time Ratio: The bottleneck time for this system is 18.34% of the total run time and it can be reduced by identifying the major bottlenecks and mitigating them.

$$\text{BN Time Ratio} = \frac{4584.5}{25000} = 0.1834$$

Bottleneck Ratio: In this system, five out of six machines are bottlenecks and so the BN ratio value is 0.83

$$\text{BN Ratio} = \frac{5}{6} = .83$$

Value Added Ratio: Since the bottleneck time ratio and bottleneck ratio are high, the value added ratio is low.

$$\text{VAR} = \frac{15831.2}{4584.4} = 3.45$$

4.4.3 Bottleneck Identification Metrics

Utilization Metric (U_V): The utilization of each machine is compared with the average utilization of the total system as shown in Table 4.13.

$$U_{avg} = \frac{75.04+71.78+58.1+32.81+78.72+63.32}{6} = 63.32$$

Table 4-13 Utilization Metric for Case Study 2

	U_1	U_2	U_3	U_4	U_5	U_6
U_V	75.04	71.78	58.1	32.81	78.72	63.32
Ratio	1.19	1.13	.91	.51	1.24	1

Since the utilization ratio of Machine 5 is large when compared to other machines, the chance of Machine 5 becoming bottleneck is large.

Queue Length Metric (W_V): The queue length of each machine is compared with the average queue length of the system as shown in Table 4.14

Table 4-14 Queue Length Metric for Case Study 2

	W_1	W_2	W_3	W_4	W_5
W_V	1.3	2	1	18.7	1
Ratio	0.27	0.41	0.21	3.89	0.21

The queue length before Machine 5 (W_4) is larger when compared to other machines and since both the bottleneck identification metrics shows the same machine as a bottleneck, it can be concluded that Machine 5 as a bottleneck.

Balancing Metric (ε_V): The degree of imbalance is high in this system as the product route is different and some machines do not process all the products. As shown in Table 4.15 machine 5 has high utilization than other machines which should be the bottleneck.

Table 4-15 Balancing Metric for Case Study 2

ε_1	ε_2	ε_3	ε_4	ε_5	ε_6
0.97316	0.94108	0.75928	0.43846	1.04267	0.84483

Variance Metric: Since there is a high imbalance between machines, the variance metric values shown in Table 4.16 cannot be considered for allotting buffers. If the number of products processed in all the machines of a system is same, the variance metric can be used for identifying the machine which needs attention through buffers.

Table 4-16 Variance Metric for Case Study 2

γ_1			γ_2			γ_3		γ_4	γ_5		
93.1	101.4	84.7	84.7	84.7	84.7	105.6	116.67	38.9	50	38.9	43.1
9.6			9.2			10.5		6.2	6.6		
γ_6											
26.4	26.4	26.4									
5.11											

Processing Time Metric (β_{SV}): In this system, there is no dominant bottleneck part as there is uniformity in the processing time metric and the values are shown in Figure 4.17.

Table 4-17 Processing Time Metric for Case Study 2

	Machine 1	Machine 2	Machine 3	Machine 4	Machine 5	Machine 6
Part 1	$\beta_{11} = 1.01$	$\beta_{12} = 1.01$	$\beta_{13} = 0.99$		$\beta_{15} = 1.02$	$\beta_{16} = 1.01$
Part 2	$\beta_{21} = 1.02$	$\beta_{22} = 0.99$			$\beta_{25} = 0.98$	$\beta_{26} = 0.99$
Part 3	$\beta_{31} = 0.97$	$\beta_{32} = 1.01$	$\beta_{33} = 1.01$	$\beta_{34} = 1.00$	$\beta_{35} = 1.00$	$\beta_{36} = 1.01$

4.4.4 Methodology

All the bottleneck identification metrics are compared to identify the bottlenecks and its type as shown in Table 4.18. In this system, Machine 5 is a dominant bottleneck machine with high queue length and imbalance values. So the improvement of the system is initially tested with buffer allocation and the capacity is added to the whole system and further it is improved by an additional capacity at bottleneck station.

Table 4-18 Comparison of Bottleneck Identification Metrics for Case Study 2

Metrics / Machines	Utilization Metric	Queue Length Metric	Balancing Metric	Variance Metric
Machine 1	1.19		0.97	9.60
Machine 2	1.13	0.27	0.94	9.20
Machine 3	0.91	0.41	0.76	10.50
Machine 4	0.51	0.21	0.44	6.20
Machine 5	1.24	3.89	1.04	6.60
Machine 6	1.00	0.21	0.84	5.11

The optimal buffer size with the maximum possible throughput is determined using the methodology mentioned in Chapter 3. The buffer sizes and throughput at various stages is shown in Table 4.19.

Table 4-19 Buffer Size and Throughput Iterations for Case Study 2

Buffer Quantity	Replications										Average Throughput
(0,0,0,0,0)	181	180	181	180	180	182	180	180	181	180	180.5
(∞,∞,∞,∞,∞)	226	224	225	225	225	225	225	225	225	225	225
(∞,∞,∞,3,∞)	225	224	223	225	225	224	225	224	225	224	224.4
(∞,∞,∞,2,∞)	207	207	208	210	208	207	207	207	207	208	207.6
(1,1,1,3,1)	225	224	224	225	225	225	225	224	226	225	224.8

(1,1,1,3,0)	223	224	223	224	223	223	225	223	222	225	223.5

Throughput for the system is improved from 180.5 to 224.4 from the buffer sizes shown in Table 4.20. Even though the buffer size is optimal, the required throughput could not be reached. Hence, an additional machine has to be added to increase throughput and the improved performance measures are shown in Table 4.21.

Table 4-20 Optimal Buffer Solution for Case Study 2

Buffer 1	Buffer 2	Buffer 3	Buffer 4	Buffer 5
1	1	1	3	1

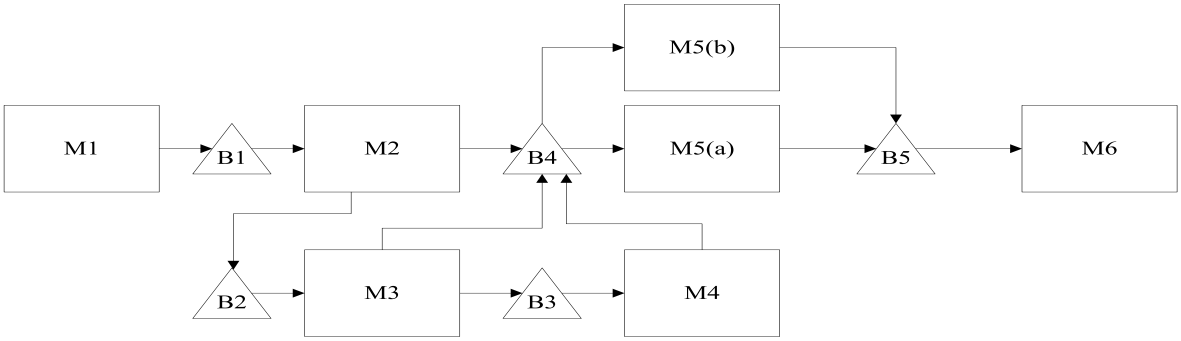

Figure 4-3 Block Diagram for Case Study 2 with Additional Capacity

Table 4-21 Performance Measures for Case Study 2

	Before	**After**
Idle time	356986.0	25926.3
Processing time	94987.1	123548.8
Blocked / Wait Time	18026.84	524.8252
Value added ratio	3.45	4.671
Throughput	180.5	224.4

The parallel machine is added based on the highest utilization technique. The new system is again analyzed for the optimal buffer solution for reaching the required throughput as shown in Table 4.22

Table 4-22 Buffer Size and Throughput with Additional Capacity for Case Study 2

Buffer Quantity	Throughput for 10 replications										Average Throughput
(∞,∞,∞,∞,∞)	241	242	240	244	240	240	239	242	239	239	240.6
(1,∞,∞,∞,∞)	241	242	242	240	240	240	240	238	238	238	239.9
(2,1,∞,∞,∞)	240	241	239	240	242	241	239	238	238	238	239.6
(2,2,1,∞,∞)	241	242	240	244	240	240	239	242	239	239	240.6
(2,2,1,10,∞)	241	242	240	244	240	240	239	242	239	239	240.6
(2,2,1,2,∞)	241	242	240	244	240	240	239	242	239	239	240.6
(2,2,1,2,3)	241	242	240	244	240	240	239	242	240	241	240.6
(2,2,1,2,2)	242	240	240	241	243	239	239	242	239	239	240.4

4.4.5 Results

The optimal buffer shown in Table 4.23 was found using the optimal buffer addition method. The required throughput which is 235 parts is reached by adding a parallel machine using the additional capacity method based on the utilization and non-value added time.

Table 4-23 Optimal Buffer for Case Study 2 with Additional Capacity

Buffer 1	Buffer 2	Buffer 3	Buffer 4	Buffer 5
2	2	1	2	3

Table 4-24 Performance Measures Comparison for Case Study 2

	Before	After
Throughput (parts)	180.5	240.9
Blocking time (min)	18026.84	19.7

4.4.6 Sensitivity Analysis

The optimal buffer is further reduced based on the economic feasibility which is discussed in economic buffer method in Chapter 3. The buffer size (2, 2, 1, 2, and 3) gives the maximum throughput and it is further reduced. As shown in Table 4.25, the buffer size is reduced based on the cost based sensitivity analysis. The buffer size (1, 1, 1, 2, and 3) is selected since the reduction in throughput is only by three parts and the cost saved from buffer space is more than the loss incurred.

Table 4-25 Economic Buffer Analysis for Case Study 2

	Buffer	**Throughput**
	(∞, ∞, ∞, ∞, ∞)	240.9
	(2,2,1,2,3)	240.9
	(1,2,1,2,3)	239
Economical Buffer	(1,1,1,2,3)	237.9
	(1,1,1,1,3)	195.8
	(1,1,1,1,2)	161

4.4.7 System Evaluation Metrics

Bottleneck Time Ratio: The bottleneck time ratio has decreased from 0.1834 to 0.159

$$\text{BN Time Ratio} = \frac{3990.1}{25000} = 0.159$$

Bottleneck Ratio: The bottleneck ratio has improved from 0.83 to 0.33

$$\text{BN Ratio} = \frac{2}{6} = .33$$

Value Added Ratio: The value added ratio has improved from 3.45 to 5.31

$$\text{VAR} = \frac{21186.39}{3990.14} = 5.31$$

4.4.8 Comparison of Different Buffer Allocation

From the results obtained, the management of a company can select any option depending on their market constraints and their requirements. Some industries which follow „make to order" policies do not need additional capacities. These different options for the above case study are shown in Table 4.26.

Table 4-26 Comparison of Different Control Strategies

	Optimal buffer	**Additional capacity**	**Economic buffer**
Throughput	225.2	240.9	237.9
Cost*	$70	$2600	$2580
Buffer	(1,1,1,3,1)	(2,2,1,2,3)	(1,1,1,2,3)

4.5 Case Study 3

A job shop type of system is considered for case study 3. In this case, 5 products are processed in 10 machines and each product has a different product route as shown in section 4.4.1. The processing time for products at each station is shown in Table 4.27.

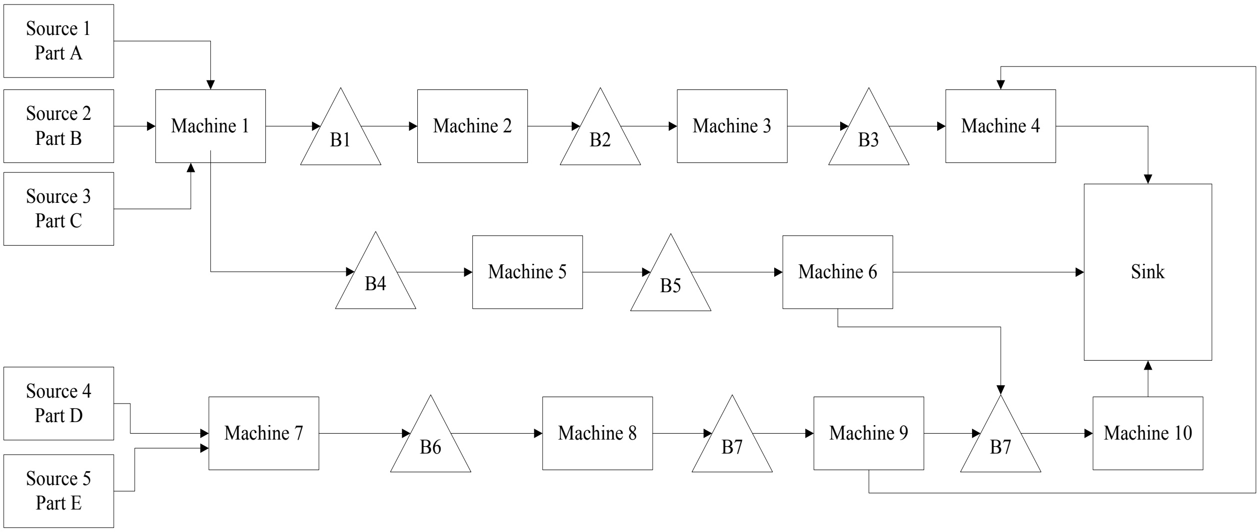

Figure 4-4 Block Diagram with 10 Machines and 5 Products for Case Study 3

4.5.1 Product Routing

The product routing and the processing time table for case study 3 is shown below.

Product 1: M1-M2-M3-M4

Product 2: M1-M2-M3-M6

Product 3: M1-M5-M6-M10

Product 4: M7-M8-M9-M10

Product 5: M7-M8-M9-M4

Table 4-27 Processing Time Table for Case Study 3

Machine 1	Part 1,2,3 – 3 Minutes			
Machine 2		Min	Mode	Max
	Part 1	2	4	6
	Part 2	3	6	7
Machine 3	Part 1	2	4	6
	Part 2	2	4	6
Machine 4	Part 1	2	6	7
	Part 5	1	8	12
Machine 5	Part 3	6	7	8
Machine 6	Part 2	3	12	17
Machine 7	Part 4	3	6	7
	Part 5	3	5	10
Machine 8	Part 4	3	7	10
	Part 5	3	6	7
Machine 9	Part 4	7	8	9
	Part 5	1	3	5
Machine 10	Part 3	5	8	11
	Part 5	5	8	11

4.5.2 System Evaluation Metrics

In this system, the bottleneck time is about 24% of the total run time. The bottleneck ratio shows that seven machines are bottlenecks at least in one instant. The value added to non value added ratio is 4.1:1. The efficiency of a system is measured using these metrics as shown in Table 4.28.

Table 4-28 System Evaluation Metrics for Case Study 3

Bottleneck Time Ratio	0.239
Bottleneck Ratio	0.7
Value Added Ratio	4.1

4.5.3 Bottleneck Identification Metrics

Utilization Metric (U_V): In this system, Machine 6 has the highest utilization as shown in Table 4.29 and the risk of Machine 6 becoming a bottleneck is high which is verified with other developed metrics.

Table 4-29 Utilization Metric for Case Study 3

U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}
0.67	0.7	0.69	1.1	0.5	1.62	1.25	1.3	1.2	0.9

Queue Length Metric (W_V): Similar to the utilization metric, the queue length metric also determines Machine 6 as the bottleneck as shown in Table 4.30. There is a significantly high queue of products before Machine 6 when compared to other machines.

Table 4-30 Queue Length Metric for Case Study 3

W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}
	0.35	0.05	0.99	0.01	5.85	0.4	0.46	0	0.28

Balancing Metric (ε_V): In this system, there is a high degree of imbalance between Machines 3, 4, 5 and Machines 9 and 10 caused by Machines 4 and 9 as shown in Table 4.31. So there are 2 machines that might have a significant impact on the system and these two machines have to be analyzed based on the other metrics in order to improve the performance of the system.

Table 4-31 Degree of imbalance for Case Study 3

ε_1	ε_2	ε_3	ε_4	ε_5	ε_6	ε_7	ε_8	ε_9	ε_{10}
1.51	1.57	1.35	2.58	1.18	1.79	2.81	2.97	2.72	3.32

Variance Metric (γ_V): In this system, Machine 6 has the highest variability and the variability is significantly higher than the other machines as shown in Table 4.32. Machine 6 which is shown as a bottleneck by utilization and queue length metric can be considered as a variance bottleneck and buffers can be allotted for improving the performance of the system. If the variability is too high, the effect of buffers would be considerably less than an additional capacity. This has to be verified by the economic analysis.

Table 4-32 Metrics Based on Variability for Case Study 3

γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}
0	0.694	0.67	3.2	0.17	8.39	1.4	1.39	0.42	1.5

4.5.4 Control Strategies

From the bottleneck identification metrics shown in Table 4.33, Machine 6 is identified as a bottleneck and since the variability is too high which also created an imbalance between the machines thereby blocking the upstream machines. So, there is a need for an additional capacity rather than additional buffers which is validated by simulation based heuristics and economic analysis.

Table 4-33 Comparison of Bottleneck Identification Metrics for Case Study 3

	Utilization Metric	Queue Length Metric	Imbalance Metric	Variance Metric
Machine 1	0.67	0	1.51	0
Machine 2	0.7	0.35	1.57	0.694
Machine 3	0.69	0.05	1.35	0.67
Machine 4	1.1	0.99	2.58	3.2
Machine 5	0.5	0.01	1.18	0.17
Machine 6	1.62	5.85	1.79	8.39
Machine 7	1.25	0.4	2.81	1.4
Machine 8	1.3	0.46	2.97	1.39
Machine 9	1.2	0	2.72	0.42
Machine 10	0.9	0.28	3.32	1.5

Additional Capacity: Machine 6 has the highest utilization, longest queue length and high variability which conclude it as a bottleneck. Since, both the queue length and utilization are high; there is a need for capacity addition. Then based on the variance metric, buffer allocation is done throughout the system. Once the capacity addition is done, the method is repeated as a process of continuous improvement. The bottleneck identification metrics after the capacity addition is shown in Table 4.34.

Table 4-34 Comparison of Bottleneck Identification Metrics after Capacity Addition

	Utilization Metric	Queue Length Metric	Imbalance Metric	Variance Metric
Machine 1	0.97	0	0.72	0
Machine 2	0.99	0.43	0.75	0.694
Machine 3	0.99	0.08	0.64	0.67
Machine 4	0.97	1.73	1.23	3.2
Machine 5	0.75	0.143	0.56	0.17
Machine 6	1.19	1.77	0.43	8.39
Machine 7	0.81	0	1.34	1.4
Machine 8	0.86	0.6	1.42	1.39
Machine 9	0.79	0.02	1.30	0.42
Machine 10	1.44	3.2	1.59	1.5

The performance of the system has been improved significantly after the capacity addition such as the 27% increase in throughput from 5063 to 6437.3 parts. For further improvement in the system throughput, the bottleneck identification metrics are used as shown in Table 4.34. The dominant bottleneck machines are 10, 6 and 4 with the highest utilization and queue length values. In such situations, the management can decide the optional control strategies between capacity addition and buffer allocation based on the economic and performance analysis as shown in Table 4.35 and 4.36. Since the throughput requirement of each product is based on demand and each machine has its own impact on different products.

Table 4-35 Buffer Allocation for Case Study 3 after Capacity Addition

Buffer 1	Buffer 2	Buffer 3	Buffer 4	Buffer 5	Buffer 6	Buffer 7	Buffer 8	Throughput
∞	∞	∞	∞	∞	∞	∞	∞	8190.5
31	9	110	1	144	51	1	153	8190.5
31	9	105	1	141	49	1	151	8150
: :	: :	: :	: :	: :	: :	: :	: :	: :
20	9	20	1	20	20	1	20	7429.6
10	9	10	1	10	10	1	10	7045.6
0	0	0	0	0	0	0	0	6437

4.5.5 Economic Analysis

The investment cost on buffers and additional capacity is compared with the performance improvement based on the throughput. The control strategy can be assumed from various options based on the customer‟s (management) requirement. The various options for buffer allocation is shown in Table 4.36

Table 4-36 Economic Analysis for Buffer Allocation

Buffer	Throughput	Cost	Profit
Infinite	8190.5		
500	8190.5	$7500	$23,770
111	7429.6	$3610	$70,686
61	7045.6	$3110	$67,346

Even though the buffer allocation improved the performance of the system, the throughput has not showed a significant improvement. The bottleneck machines which had a high utilization and queue length along with high imbalance metric needs additional capacity to eliminate the bottlenecking effects. Once the additional capacity is added, the performance of the system has increased significantly as shown in Table 4.37.

Table 4-37 Economic Analysis for Capacity Addition

Buffer	Throughput	Cost	Profit
0	6437	$2500	$61,870
0	5063	$0	0
31	10075	$7810	$92,940

4.5.6 Results

After the control strategies are implemented based on the bottleneck identification metrics and the heuristics, the efficiency of the system has improved significantly as shown in Table 4.38. Even though there are many options for selecting an optimal strategy, the system efficiency metrics are calculated based on the lowest investment cost.

Table 4-38 System Efficiency Metrics

	Before	After
Bottleneck Time Ratio	0.239	0.11
Bottleneck Ratio	0.7	0.2
Value Added Ratio	4.1	24.47
Throughput	5063	7429

The options such as optimal buffer, additional capacity and economical buffer is shown in Table 4.39. Based on the requirement, available layout space and available investment, any option can be selected.

Table 4-39 Comparison of Control Strategies

	Optimal buffer	Additional capacity	Economic buffer
Throughput	8190.5	10,075	7045.6
Cost*	\$7,500	\$7810	\$3,110
Buffer	500	31	62

*Cost/buffer space = \$10, Cost/capacity = \$2500

CHAPTER 5

CONCLUSION AND FUTURE RESEARCH

In this chapter, the conclusion and the possible future research from this thesis is detailed. The methods for identifying bottlenecks in complex production system and mitigating techniques are discussed in Chapter 3. Simulation based heuristics for mitigating bottlenecks such as buffer allocation, capacity addition, and economic analysis are developed and explained in the Chapter 3. Further, the various scenarios which are possible in complex production system are listed and case studies are developed and solved in Chapter 4. In this chapter, Section 5.1 discusses about the bottleneck identification method and Section 5.2 discusses about the bottleneck mitigation strategies. Section 5.3 discusses about the possible extensions that can be considered for future research.

5.1 Bottleneck Identification

The first stage of this research is identifying the bottlenecks in complex production systems. Bottlenecks in this system are identified using various metrics such as utilization metric, queue length metric, variance metric, and balancing metric. Even though the utilization and queue length based metrics are proposed earlier by other researchers, they were not used as a combination. In this research, all these metrics are integrated to identify bottleneck accurately even in complicated systems. The variance metric and imbalance metric would determine the type of bottleneck involved in the system. Finally the processing time metric, is used to find the particular product which makes the machine bottleneck.

5.2 Bottleneck Mitigation

In this research, bottleneck mitigation has three-fold objective. Based on the values obtained from metric values, the type of bottleneck is determined and the machine which needs most attention is assessed. The machines with variability without an imbalance can be mitigated using buffers to an extent. If the variability is very high where adding a capacity would have a significant impact than adding buffers, decision is made based on the economic analysis. Simulation based heuristics are developed for buffer allocation based on the qualitative characteristics such as continuity, monotonicity, concavity and discontinuity. The additional capacity is added to the machine which has highest utilization, longest queue length, and high imbalance. Based on these metric values, the machine which needed an additional capacity is mitigated by the machine addition heuristic shown in Chapter 3. Once the performance of the system is improved by buffer allocation and capacity addition for meeting the throughput, an economic analysis is done for analyzing the feasibility of mitigation. Since the most recognized technique of performance enhancement is continuous improvement, these methods are also continuously used for mitigating various bottlenecks as there is always a bottleneck process in any system.

5.3 Future Research

Though this research concentrated on complex production systems; there are some research extensions that can be done to further improve the mitigation of bottlenecks. A common buffer location can be installed in the system, where parts from multiple machines can be stored at a common location. The common buffer location is more economical as the buffer space before each machine would be idle if the machines in the system have micro-breakdowns. The location of common buffer storage can be determined using the material handling costs based on the flow of products.

Another opportunity in improving the reliability of the system is by allocating instantaneous buffers to the system. In this method, the buffers are allocated based on the events occurring in the system. A live simulation model can be developed along with a program for automatic buffer allocation which updates the system events. For example, if a machine fails, the buffers are allocated at that particular instant in order to have a smooth flow of products throughout the system. This method can reduce the work in progress and the excess space allocated for buffers considerably and can be integrated with the common buffer storage method. The buffer allocation method can assume on market constraints which would consider more realistic scenarios.

REFERENCES

LIST OF REFERENCES

Altiparmak, F., Dengiz, B., and Bulgak, A. A., (2007), “Buffer Allocation and Performance Modeling in Asynchronous Assembly System Operations: an Artificial Neural Network Meta Modeling Approach”, *Applied Soft Computing*, Vol. 7, pp. 946–956.

Anderson, D. R. and Moodie, C.L., (1969), “Optimal buffer storage capacity in production line systems”, *The International Journal of Production Research*, Vol. 7, No. 3, pp. 233-240.

Battini, D., Persona, A., and Regattieri, A., (2009), “Buffer Size Design Linked tTo Reliability Performance: A Simulative Study”, *Computers & Industrial Engineering*, Vol. 56, pp. 1633–1641.

Boysen, N., Malte Fliedner, M., and Scholl, A., (2007) “A Classification of Assembly Line Balancing Problems”, *European Journal of Operational Research*, Vol. 183, pp. 674–693.

Bulgak, A. A., Diwan, P. D., and Inozu, B., (1995), “Buffer Size Optimization in Synchronous Assembly Systems Using Genetic Algorithms”, *Computers in Industrial Engineering*, Vol. 28. No. 2, pp. 309-322.

Chiadamrong, N. and Limpasontipong, P., (2003), Using Storage Buffer to Improve Unbalanced Asynchronous Production Flow Line‟s Performance, *International Journal for Manufacturing Technology and Management*, Vol. 5, pp. 149-161.

Chiang, S. Y., Kuo, C. T., and Meerkov, S. M., (1998), “Bottlenecks in markovian production lines: A systems approach,” *IEEE Transactions on Robotics and Automation*, Vol. 14, No. 2, pp. 352-359.

Chiang, S. Y., Kuo, C. T., and Meerkov, S. M., (2002), “C-bottlenecks in serial production lines: Identification and application,” *Mathematical Problems in Engineering*, Vol. 7, pp. 543-578.

Cochran, J. K. and Kokangul, A., (2009),”Stochastic Approximations for Optimal Buffer Capacity of Many-Station Production Lines”, *International Journal for Mathematics in Operational Research,* Vol. 1, pp. 211-227.

Conway, R., Maxwell, W., McClain, and Thomas, J., (1988), “The Role of Work-In-Process Inventory In Serial Production Lines”, *Operations Research*, Vol. 36, No. 2, pp. 229-241.

Faria, J., Matos, M., and Nunes, E., (2006), “Optimal Design of Work-In-Process Buffers” *International Journal of Production Economics,* Vol. 99, pp. 144–155.

Gardiner, S. C. And Blackstone, J. H. (1998), “Dynamic buffering”, *International Journal of Production Research*, Vol. 36, No. 2, pp. 333-342.

Gershwin, S. B. and Schor, J. E., (2000), “Efficient Algorithms for Buffer Space Allocation”, *Annals of Operations Research*, Vol 93, pp. 117–144.

Harris, J.H. and Powell S. G., (1999), “An Algorithm for Optimal Buffer Placement in Reliable Serial Lines”, *IIE Transactions*, Vol. 31, pp. 287-302.

Hillier, F.S. and Boling, R. W., (1966), “The Effect of Some Design Factors on the Efficiency of Production Lines with Variable Operations Times”, *Journal of Industrial Engineering*, vol. 17, pp. 651-658.

Hillier, F.S., So, K.C., and Boling, R. W., (1993), “Toward Characterizing the Optimal Allocation of Storage Space in Production Line Systems with Variable Processing Times”, *Management Science*, Vol. 39, No. 1, pp. 126-133.

Hillier, M. S. And Hillier, F. S., (2006) “Simultaneous Optimization of Work and Buffer Space in Unpaced Production Lines with Random Processing Times”, *IIE Transactions*, Vol. 38, pp. 39–51.

Hillier, M.S., (2000), “Characterizing the Optimal Allocation of Storage Space in Production Line Systems with Variable Processing Times”, *IIE Transactions*, Vol. 32, pp. 1-8.

Huang, M. G., Chang, P. L., and Chou, Y., (2002), “Buffer Allocation in Flow-Shop-Type Production Systems with General Arrival and Service Patterns”, *Computers & Operations Research*, Vol. 29, pp. 103-121.

Jeong, K.C. and Kim, Y. D., (2000), “Heuristics for Selecting Machines and Determining Buffer Capacities in Assembly Systems”, *Computers and Industrial Engineering*, Vol. 38, pp. 341-360.

Kouikoglou, V. S. and Phillis, Y. A., (1997), “A continuous flow model for production networks with finite buffers, unreliable machines, and multiple products”, *International Journal of Production Research*, Vol. 35, NO. 2, 381-397.

Kuo, C. T., Lim, J. T., and Meerkov, S. M., (1996), “Bottlenecks in serial production lines: A system-theoretic approach,” *Mathematical Problems in Engineering*, Vol. 2, pp. 233-276.

Law, A. M., and Kelton, D. W., (2000), *Simulation modeling & analysis,* Mcgraw Hill.

Lawrence, S. R., and Buss, A. H., (1994), “Shifting production bottlenecks: Causes, cures and conundrums,” *Journal of Production and Operations Management*, Vol. 3, No. 1, pp. 21-37.

Lawrence, S. R., and Buss, A. H., (1995), “Economic analysis of production bottlenecks,” *Mathematical Problems in Engineering*, Vol. 1, No. 4, pp. 341-369.

Li, L., Chang, Q., and Ni, J., (2008), “Data driven bottleneck detection of manufacturing systems,” *International Journal of Production Research*, Vol. 47, No. 18, pp. 5019-5036.

Lutz, C. M., Davis, K.R., and Sun, M., (1998), “Determining Buffer Location and Size in Production Lines Using Tabu Search” *European Journal of Operational Research*, Vol. 106, pp. 301-316.

Meller, R. D. and Kim, D.S., (1996), “The impact of preventive maintenance on system cost and buffer size”, *European Journal of Operational Research*, Vol. 95 pp. 577-591.

Nahas, N., Ait-Kadi, D., and Nourelfath, M., (2006), “A New Approach for Buffer Allocation in Unreliable Production Lines”, *International Journal of Production Economics*, Vol. 103, pp. 873-881.

Nahas, N., Nourelfath, M., and Ait-Kadi, D., (2009), “Selecting Machines and Buffers in Unreliable Series-Parallel Production Lines”, *International Journal of Production Research*, Vol. 47, No. 14, pp. 3741–3774.

Nieuwenhuyse, I. V., Vandaele, N., Rajaram, K., and Karmarkar, U.S., (2007), “Buffer Sizing in Multi-Product Multi-Reactor Batch Processes: Impact of Allocation and Campaign Sizing Policies”, *European Journal of Operational Research*, Vol. 179, pp. 424–443.

Othman, Z., Kamaruddin, S., and Ismail, M. S., (2007), “Optimal Buffer Allocation for Unpaced Balanced and Unbalanced Mean Processing Time”, *Jurnal Teknologi*, Vol. 46, pp. 31–42.

Papadopoulos, H.T. and Vidalis, M.I., (2001), A Heuristic Algorithm for the Buffer Allocation in Unreliable Unbalanced Production Line, *Computers and Industrial Engineering,* Vol. 41, pp. 261-277.

Patti A. L., Watson, K., and Blackstone, J. H., (2008), The Shape of Protective Capacity in Unbalanced Production Systems with Unplanned Machine Downtime, *Production Planning and Control*, Vol. 19, No. 5, pp. 486-494.

Powell, S.G. and Pyke, D.F., (1998), “Buffering Unbalanced Assembly Systems”, *IIE Transactions*, Vol. 30, pp. 55-65.

Qudeiri, J. A., Yamamoto, H., Ramli, R. and Jamali, A., (2008), “Genetic Algorithm for Buffer Size and Work Station Capacity in Serial–Parallel Production Lines”, *Artif Life Robotics*, Vol. 12, pp. 102–106.

Radovilsky, Z.D., (1998), “A Quantitative Approach to Estimate the Size of the Time Buffer in the Theory of Constraints”, *International Journal of Production Economics*, Vol. 55, pp. 113-119.

Ribeiro, M. A., J.L. Silveira, J.L., and Qassim, R.Y., (2007), “Joint Optimisation of Maintenance and Buffer Size in a Manufacturing System”, *European Journal of Operational Research*, Vol. 176, pp. 405–413.

Roser, C., Nakano, M., and Tanaka, M., (2001), “A practical bottleneck detection method,” *Proceedings of the Winter Simulation Conference*, pp.949-953.

Roser, C., Nakano, M., and Tanaka, M., (2002), "Shifting bottleneck detection," *Proceedings of the Winter Simulation Conference*, pp. 1079-1086.

Roser, C., Nakano, M., and Tanaka, M., (2003), "Buffer allocation model based on a single simulation," *Proceedings of the Winter Simulation Conference*, pp. 1238-1246.

Sabuncuoglu, I., Erel, E., and Gocgun, Y., (2006), "Analysis of Serial Production Lines: Characterisation Study and a New Heuristic Procedure for Optimal Buffer Allocation", *International Journal of Production Research*, Vol. 44, No. 13, pp. 2499–2523.

Shi, C. and Gershwin, S. B., (2009), "An Efficient Buffer Design Algorithm for Production Line Profit Maximization", *International Journal of Production Economics*, Vol. 122, pp. 725–740.

Shi, L. and Men, S., (2003), "Optimal Buffer Allocation in Production Lines", *IIE Transactions*, Vol. 35, pp. 1–10.

So, K.C., (1997), "Optimal buffer allocation strategy for minimizing work-in-process inventory in unpaced production lines", *lIE Transactions,* Vol. 29, pp. 81-88.

Tamilselvan, P., Krishnan, K., and Cheraghi, H., (2010), "Measurement of shifting bottlenecks in a non-buffer production system", *Proceedings of the Industrial Engineering Research Conference*, Cancun Mexico, June 2010

Vergara, H. A. and Kim, D. S., (2009), "A New Method for the Placement of Buffers in Serial Production Lines", *International Journal of Production Research*, Vol. 47, No. 16, pp. 4437–4456.

Yamashita, H. and Altiok, T., (1998), "Buffer Capacity Allocation for a Desired Throughput in Production Lines", *IIE Transactions,* Vol. 30, pp. 883-891.

Ye, T. and Han, W., (2006), A quantitative method to determine the size of the stock buffer in front of the bottleneck under multi product, *Proceedings of the Sixth World Congress on Control and Automation*, Vol. 2, pp. 7239-7243.