2011

An Analytical Approach to Cycle Time Evaluation in an Unreliable Multi-Product Production Line with Finite Buffers

Farshad Jarrahi
University of Windsor

An Analytical Approach to Cycle Time Evaluation in an Unreliable Multi-Product Production Line with Finite Buffers

by

Farshad Jarrahi

A Thesis
Submitted to the Faculty of Graduate Studies
through Industrial and Manufacturing Systems Engineering
in Partial Fulfillment of the Requirements for
the Degree of Master of Science at the
University of Windsor

Windsor, Ontario, Canada

2011

Your file Votre référence
ISBN: 978-0-494-80225-0
Our file Notre référence
ISBN: 978-0-494-80225-0

DECLARATION OF ORIGINALITY

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

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

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

ABSTRACT

This thesis develops an analytical approximation method to measure the performance of a multi-product unreliable production line with finite buffers between workstations. The performance measure used in this thesis is Total Cycle Time. The proposed approximation method generalizes the processing times to relax the variation of product types in a multi-product system.

A decomposition method is then employed to approximate the production rate of a multi-product production line. The decomposition method considers generally distributed processing times as well as random failure and repair. A GI/G/1/N queuing model is also applied to obtain parameters such as blocking and starving probabilities that are needed for the approximation procedure.

Several numerical experiments under different scenarios are performed, and results are validated by simulation models in order to assess the accuracy and strength of the approximation method. Consequent analysis and discussion of the results is also presented.

DEDICATION

I dedicate this thesis

To my beloved mother,

and

To honor the memory of my late father

ACKNOWLEDGEMENTS

This masters study and research has provided me the opportunity to enhance my knowledge in the area of Industrial Engineering. I would like to express my earnest gratitude to all those who gave me the opportunity to complete this thesis.

I am deeply indebted to my supervisor Dr. Walid Abdul-Kader whose leadership, cooperation, insights, and encouragements have always enlightened and directed me toward achieving my academic goals. He helped me in all the time of this research and writing of this thesis. I want to thank the committee members Dr. Shervin Erfani, and Dr. Fazle Baki for their constructive criticism and valuable suggestions regarding this research work. I would also like to thank Dr. Tirupati Bolisetti for chairing me thesis defence.

I must acknowledge the following staff members – Mrs. Angela Haskell, Ms. Qin Tu, Mrs. Jacquie Mummery, Mrs. Brenda Schreiber, Mr. Dave McKenzie, Mr. Ram Barakat in the IMSE department, Mrs. Svetlana Georgieva in the Graduate Studies Office, and my friend Saman Hassanzadeh for being helpful at all levels.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

A manufacturing production line is a common type of production system. It is also known as transfer line or flow line and can be represented as a finite buffer tandem queuing system. In this thesis, we use the terms "production line" and "flow line" interchangeably. A production line consists of a number of workstations or machines in series with intermediate buffers between successive workstations or machines. Examples of flow lines can be observed in electronic components assembly and high-volume automotive parts production industries.

In a synchronized flow line system such as automatic transfer line and automatic assembly line, materials transfer from one work centre to the next at the same instant. In this system, the machines in the line are highly coupled and as soon as any machine breaks down all other machines in the line will be forced down at the next transfer instant. Unlike synchronized system, in an asynchronized flow line system, transfer of materials from one work centre to the next begins as soon as its processing at a work centre is complete and the next work centre has become free. In this system, a machine failure is considered to be equivalent to an extended processing time and eventually causes all other machines in the line to be forced down.

The amount of time the material (part) spends in a workstation may be considered as deterministic if it does not vary from one part to the next ,and stochastic if it varies randomly from part to part. This randomness may be due to random processing times, random failure and repair events, or both (Dallery and Gershwin, 1992). The production line dealt with in this thesis is asynchronous in nature, and failures and repairs are random with operation dependent failures. Parts are transferred to the next workstation via the intermediate buffer on a unit-by-unit basis.

In a flow line system, intermediate buffers are used to decouple the machines and to provide them with some independence in their operations. When a machine in a flow line fails or has a longer operation time, the machine upstream of it can still operate until the upstream buffer fills up, and the downstream machine can still operate until the downstream buffer becomes empty. If the buffer downstream of a machine fills up, the machine has to temporarily stop production until the buffer has space for more output. In this case, the machine is said to be blocked. A machine may also have to stop temporarily if the buffer upstream of it is empty; now the machine is said to be starved. The introduction of buffer minimizes the blocking and starvation times of the machines.

There are different ways to improve the performance of a flow line system, such as:

(a) Increase intermediate storage capacity (Buzacott, 1971), because such buffers may allow the machines to continue working when one of them is down

(b) Improve the availability of machines by providing some degree of parallelism at the stage level, because parallel machines can continue the function when one or more (not all) of the working machines fails (Aziz et al. 2010)

(c) Increase the working capacity of machines

(d) Perform preventive maintenance of the machines to improve machine reliability and hence availability by reducing failure rate

Every solution described above has merits as well as limitations. An increase of intermediate storage capacity and addition of parallel machines could be expensive and requires valuable floor space at the plant. There are also other cost factors involved such as holding cost and material handling cost associated with the in-process inventories on the buffer. Increasing the working capacity of machines is restricted by the design of the machines. Every machine has its designed maximum and rated speed. Running a machine beyond this recommended speed decreases the reliability and eventually the life of the machine. Failures of the machines may not be completely eliminated in a manufacturing system. However, they can be reduced with the implementation of preventive maintenance in a planned manner. The benefit of preventive maintenance depends on the frequency of use and costs involved. So, for the selection among the above solutions a performance evaluation model that shows the economic tradeoffs is needed.

There exist many different metrics that can be applied as a measure of performance in a manufacturing system. Production rate, average number of parts stored in intermediate buffers, and throughput time are among the most commonly used measures of performance in the literature. In this thesis we will introduce and use Total Cycle Time as the ultimate performance measure (the definition will be given later in this thesis).

Today, with the help of low cost automated machines and material handling systems, many small and medium sized companies are equipped with automated serial production systems capable of producing many different product types without major changes in the system configuration. A good example of such systems is packaging lines in pharmaceutical or food industries that produce a large variety of product types.

However, in such systems, analyzing the performance will become complicated due to the variety of product types. Most of the existing studies in the area of manufacturing performance measurement have been focused on classic production lines that are traditionally assumed to produce single product type. In this thesis, we will present an approach to tackle performance evaluation problem in production lines that produce more than one product types.

In chapter 2 of this thesis, a review of literature is presented based on modeling techniques. Chapter 3 introduces notation, description of the problem, assumptions and develops a methodology to approximate Total Cycle time as the ultimate performance measure. It tackles the complexity of product type variation by introducing general performance measures that are independent of product type changes. It employs GI/G/1/N queuing model and decomposition method to approximate the general production rate as a key performance measure in this thesis.

Chapter 4 presents the application of the proposed analytical approximation method on several cases to examine its accuracy, strength and limitations. The method is compared with the results obtained by simulation. Finally, conclusions and future research directions are discussed in chapter 5.

CHAPTER 2

LITERATURE REVIEW

In this chapter, a review of previous research works in the area of performance evaluation of manufacturing systems is presented. This review is organized based on model development techniques. We can categorize models by their computational form. With this taxonomy, models are analytical or experimental (Askin, G., and Standridge, R., 1993). Analytical models represent a more mathematical abstraction of the real system and are mainly divided into two major sub-categories: Exact models and approximation models. Simulation models, on the other hand, are experimental. In addition to these categories, hybrid models may also be used. More explanation of each model development method and the related literature review is given in this chapter.

2.1 Exact Models

Johri (1987) developed a Linear Programming model in order to measure the cycle time in an automated serial production line where each workstation is prone to failure. The developed LP model provides a tool for the decision maker to study and analyze the effect of factors such as buffer size, batch (job) sequence, and batch size on the production line performance. However, stochastic nature of production systems is not considered in this study. Processing times are assumed to be deterministic and workstations are reliable (no failure can happen).

Buzacott (1971) presented some simple models of production-inventory systems and demonstrated how production capacity and flexibility is enhanced with the use of inventory banks. The models were developed with different sets of assumptions of breakdowns, processing times, repair times, and times between failures, and then, the usefulness and benefits of inventory banks (buffers) were analyzed. Unfortunately, as like as many other researchers in this area this study is limited to a single-product manufacturing system.

A linear programming model was proposed by Abdul-Kader (2006) by exploiting Johri (1987)'s multi-product model and further modified it by replacing machines' repair and downtimes with the insertion of fictive products to address the issue of capacity estimation/ improvement of a multi-stage, unreliable serial production line with finite buffers. The line could process a variety of products in batch according to a predefined sequence. Buffer contribution to minimize the cycle time of the production line was also addressed through experimental optimization. Compared to the Markovian approach and the approximation methods, this approach considers more parameters that have a direct impact on the capacity of the production line. These are the variety of products, the set-up times. Longer lines can be solved in a reasonable period of computational time. However, only deterministic processing times are considered in the proposed model.

Gaver (1962) considered a single server queue with Poisson arrival and general independent service time where the server is subject to random failure, and repair time is generally and independently distributed. This article introduces the notion of completion time as the duration of the period that elapses between the instant at which the service of an Arrival (of the nth part) begins and that at which process of the next (n+1)th does begin. Using this definition, it studies the effect of a certain class of interruptions (failures and repairs) upon a waiting line process. However, in this study, only a single-server single phase queuing model is analyzed.

Zipkin (1995) developed a performance measuring method for a flexible production-inventory system in a multi product environment. The whole production facility is assumed as a single server M/G/1 queue. And then, two control policies are defined: First-come-first-serve (FCFS) policy, and Longest Queue (LQ) policy. A closed form measure of performance is derived for the first policy (FCFS), and an approximation method is proposed for the LQ policy. Finally, the numerical results are compared against the exact solution and simulation.

Banik and Gupta (2006) carry out the analysis of a finite and infinite buffer GI/MSP/1 queue (MSP stands for Morkovian Service Process; further explanation of queuing theories is given in chapter 3 of this thesis). For the finite buffer, they used the method of supplementary variable and embedded Markov chain. They also investigate the corresponding infinite capacity queuing system using a combination of the matrix-geometric technique and the classical argument based on the key renewal theorem. They derive explicit analytic expressions for the steady-state system length distribution at pre arrival and arbitrary epochs. The analysis of actual waiting time (in the system) is also presented.

Albores–Velasco and Tajonar-Sanabria (2004) analyze a queuing system with c servers, recurrent general arrivals, service type of Markovian type and common buffer of finite capacity: GI/MSP/C/R. They continued the work of Bocharov (1996) who studied such a system but for one channel (server) case. Their study is based on a multi-dimensional ordering of the state set of the system.

Asmussen and Moller (2001) study waiting time distribution in GI/PH/c and MAP/PH/c (MAP stands for Markovian Arrival Process and PH for Phase type distribution) queues. They evaluate the steady state distribution of the waiting time W in a many-server queue with servers each having a phase-type service time distribution. First, they establish their model to tackle a queuing system where the service time distribution of servers is heterogeneous. Then, they develop their model in order to analyze the cases of service time being homogeneous. They derive the phase-type representation in a form which is explicit up to the solution of a matrix fixed point problem. They develop a model to calculate the steady state density and waiting time distribution for heterogeneous, and then homogenous servers. Their key new ingredient is a study of the not-all-busy period where some or all servers are idle.

Xu and Zhang (2006) investigate a multi-server with single vacation policy. Their model is suitable for a manufacturing system in which preventive maintenance is performed during servers' idle time. They consider a quasi-birth-and-death (QBD) process in order to characterize their model and to develop various stationary performance measures for their system.

Unfortunately, this author as like as most of other researchers in the area of queuing theories has considered multi servers only in parallel, and not in series.

Alexandros & Chrissoleon (2008) propose an exact solution to a two stage one buffer flow line in which each stage consists of parallel servers. A Markov process model was analyzed and the transition equations of their model were derived and an algorithm that generates the transition matrix for any buffer size was developed. Once the transition equations are solved the performance measures for their model can be evaluated.

Exact mathematical methods usually provide us with optimal solutions and are analyzed using basic formulae and Markov processes, and in general via the use of transform methods. However, results are difficult to obtain due to the mathematical complexities, and are only available for short production lines. In theory, systems can be modeled via Markov chains for any number of stages; but in practice, it is very difficult to obtain exact analytical solutions of transfer lines with more than two machines. The reason is the number of system states in the Markov chain increases exponentially with the increase of machines and the inter-stage buffer capacity. For example, a line with four machines and inter-stage buffers of capacity 3 gives rise to a Markov chain with 19,402 states (Hillier and So, 1991).

Therefore, to cope with the complexity of exact mathematical solutions we make use of approximation methods; they attempt to use a rational method to find a good (nearly optimal) solution to the problem. In continue, a review of some existing approximation methods is given.

2.2 Approximation Models

Gershwin (1987a) extended the work of Schick and Gershwin (1978) on two-stage model using decomposition technique. The idea of the decomposition technique is to decompose the analysis of a multi-stage line into the analysis of a set of two-machine lines, which are much easier to analyze. The set of two-machine lines is assumed to have equivalent or similar behaviors to the original system. He assumed a homogeneous processing rate but different repair and failure rates for all the machines in the line. Gershwin's model demonstrated the tradeoffs between buffer capacity and throughput. The algorithm was slow, lacked robustness and failed to converge at times.

The efficiency of Gershwin (1987a) algorithm was improved upon by Dallery et al. (1988) who proposed an efficient algorithm, known as DDX algorithm that solved the Gershwin equations in substantially less time.

Gershwin (1987b) made an attempt to treat three-parameter continuous material systems where he combined two equal processing rate machines together. The concept was, one machine would be accountable for the failure rate and the other would modify the machine speed. But this idea was proved unrealistic after the simulation experiments.

Tempelmeier and Burger (2001) propose a decomposition method in order to analyze the performance of non-homogeneous asynchronous Flow Production Systems (FPSs). They consider characteristics such as general processing time, unreliable stations, and scrapping for the studied FPSs. A GI/G/1/N queuing model is also used as a part of their proposed decomposition method. They compare their numerical results with an existing exact model for small reliable lines and with simulation results for larger unreliable lines. The numerical study indicates that the approximation quality increases when the possibility of simultaneous blocking and starving is explicitly considered. It also shows a good accuracy of the proposed method especially for situations where the buffer size is in the same order of magnitude.

De Koster (1987) used repeated aggregation to multi-stage continuous flow lines of several unreliable machines separated by buffers for the prediction of line efficiency. He considered time dependent failure of the machines with exponentially distributed life and repair times. The basic idea of the aggregation technique is to reduce the system dimension by replacing a two-machine-one-buffer sub-line by one single equivalent machine in the system. Then this equivalent machine is combined with a buffer and a machine of the original line to form a new two-machine one-buffer sub-line, which is then aggregated into a single equivalent machine. This process is repeated until the last or first machine is reached, depending on the direction the aggregation is performed (Dallery and Gershwin 1992).

Haskose et al. (2003) develop an approximation model and algorithm for finding analytically the steady-state solutions for any form of arbitrary queuing networks. They apply work load control ideas to manage the operation of Make to Order production systems. They present various sets of experiments to examine and analyze relative value of work load control for controlling manufacturing lead times, job release and order acceptance, and to gain insight into how increased complexity can affect the performance measure.

In spite of the capability of solving considerably longer production lines, unfortunately, none of the aforementioned approximation methods do consider a production system with more than one product type.

2.3 Simulation Models

Methodologies other than Markov chain analysis and approximation techniques have also been used for model development. Simulation models are experimental and mimic the events that occur in the real system, allowing experimentation with operating parameters or control logic (Askin and Standridge, 1993).

Conway et al. (1988) used simulation to investigate the behaviour of serial buffered lines due to lack of synchronization and explored the distribution and accumulation of work-in-process (WIP) inventory. Based on the results obtained, they presented rules about the optimal buffer allocation/ location.

Powell and Pyke (1996) presented a simulation based approach to analyze reliable and asynchronous serial lines with variable processing times. They focused on the allocation of optimum buffer in the production lines with bottleneck machine and showed how the optimal allocation depends on the location and severity of the bottleneck, imbalances in the mean processing times, length of the line, as well as the number of buffers available.

Abdul-Kader and Gharbi (2002) developed an experimental design approach based on a simulation model to evaluate the capacity of a multi-product multi-stage unreliable production line. The model considered explicitly the variety of products, set-up times as product type changes, and the failure and repair of workstations. An experimental design was used to locate/allocate buffer spaces to assess their effectiveness and impact on the overall capacity of the production line.

In spite of the advantages of using simulation for analysis of performance such as high flexibility of modeling in terms of studying behavior of system under different scenarios, simulation modeling and analysis can be time consuming and expensive. Especially, where a problem can be solved analytically, use of simulation is not suggested (Banks et al., 2005).

2.4 Hybrid Models

The hybrid models combine different methods to evaluate the performance of production systems. A simulation model of a complex system could be built where modules of the real system are replaced by simple analytic models. The logic of the simulation model relates the individual modules together to replicate overall system occurrences.

Blumenfeld (1990) combined the result of prior theoretical and simulation studies into an approximate expression for the production rate of serial lines with identical random processing time distributions at each stage and identical buffer capacities between each pair of stages. Baker et al. (1994) adopted the same technique to predict the throughput of unbalanced three-machine serial lines without intermediate buffer.

Kenne and Gharbi (2004) present a hybrid model combining an analytical approach, simulation, and experimental design in order to choose production rates that minimize the expected cost of inventor/backlog. The system under study consists of one unreliable machine producing two part types. At first, they propose an analytical model to develop a mathematical representation of the system based on some simplified hypothesis. This model is parameterized by the thresholds of production rates as control policy. Then, they propose a simulation model that consists of the data input module, the simulation module and the output analysis module. The inputs of the simulation model are production rate parameters, and the output is incurred cost. Finally a response surface experiment design is conducted to determine which input factors or their interactions have significant impact on the incurred costs. From this estimated relation, the optimal values of the estimated input factors are determined.

The following three models are cited due to their potential use in a future research and extension to this thesis.

Choudhury (1998) formulates production systems with batch arrival of raw materials as a $M^X/M/1$ queue with random setup and vacation time. In their model, a multi vacation policy in which a server is turned off when the queue is empty for a random amount of time and stays in vacation until first customer's arrival. The arrival time of batches are Poisson, but the set up time and vacation time (while the server is turned off) are generally distributed. However, the model discussed in this paper is not exactly the same as usual vacation model but its analysis is almost similar to those in vacation model.

Ke (2007) studied an M/G/1 queuing system under vacation policies in order to analyze batch arrival systems. The author investigated their model under two vacation policies: (1) in a single queue single server model when there is no customer in the system, the server shuts down its service and goes to a vacation for a random amount of time (which has general probability distribution function). When the server returns from the vacation, if there is at least one customer in the line it starts to work. When there is no customer in line, there are two policies: (1) multiple vacation policy in which the server goes to another vacation and this iteration repeats until there is at least one customer in the line, or (2) single vacation policy in which the server waits until the first customer arrives and then it starts to work. A closedown time is needed before the server is deactivated and a start-up time is needed before the server is reactivated. Considering the server is unreliable, vacation time, service time and repair time are generally distributed, and arrival time and time to failure are exponentially distributed. This investigation is a reflection of more practical environments and can be applied as performance measures in batch production manufacturing or workshop systems.

Gupta and Banik (2007) analyzed finite buffer single server queue when arrivals occur in batches of random size. They assumed that inter arrival times are generally distributed and service process is correlated, and is represented by Markovian service process (MSP). Therefore, they considered a GI^X /MSP/1/N queue. Their study includes two batch rejection policies; 1- Total batch rejection (batches that upon arrival do not find enough space in the buffer are fully rejected), and 2- Partial batch rejection policy. Under these two circumstances, they analyzed steady distribution at pre-arrival and arbitrary epochs. Finally, the performance measures of this queuing model are calculated and demonstrated through a numerical process.

Based on the research papers reviewed, there are few known studies that evaluate serial production lines manufacturing more than one product type. This thesis introduces a new approach in modeling and evaluating the performance of manufacturing systems with multiple products. We address this issue for a considerably longer production line system with finite buffer which is common in industry. We consider general and/or deterministic processing times, operation dependent failures of the machines with exponentially distributed failures and repairs. Our model can employ more than one product types with predefined sequence and lot size. We also consider set-up times of the machines involved as the product type changes.

CHAPTER 3

RESEARCH PROBLEM AND SOLUTION METHODOLOGY

3.1 Introduction

In the previous chapter, a review of literature in the area of the performance evaluation of flow lines has been presented. This chapter introduces the statement of research problem, notation and definition of terms, assumptions, and the solution methodology that we apply and deal with in this thesis.

3.2 Problem statement

This thesis addresses the problem of performance evaluation in a production line that produces several different product types with predefined sequence and sizes. Machines are unreliable due to random failure, and set up for the whole line is needed when product type changes. This problem will be stochastic in nature due to random failure and repair rates.

The targeted performance measure is Total Cycle Time; that is the total time needed for the production of all product types. Therefore, the problem we are dealing with is finding out the Total Cycle Time. Also, it will be shown latter that we need to calculate several other measures in order to obtain Total Cycle Time, and the most important of them is Production Rate (X). In this thesis, an approximation method will be proposed to calculate Total Cycle Time as a measure of performance in a multi-product production line.

The schematic diagram presented in Figure (3.1) illustrates the model to be studied which is a production line with M workstations buffered with (M−1) buffers. The flow direction is indicated by the arrows. As demonstrated, the aforementioned production line processes different product types without major changes in the system configuration (a good example of such systems can be packaging lines). As soon as the first part (piece) of each batch arrives at the first workstation of the line, a set up is needed for the whole line before the production starts. The reason that we consider set up for the whole line, instead of each workstation individually, is the interaction and dependency that might exist between the set up of adjacent machines; particularly, in situations where a production line uses automated material handling system (such as conveyor).

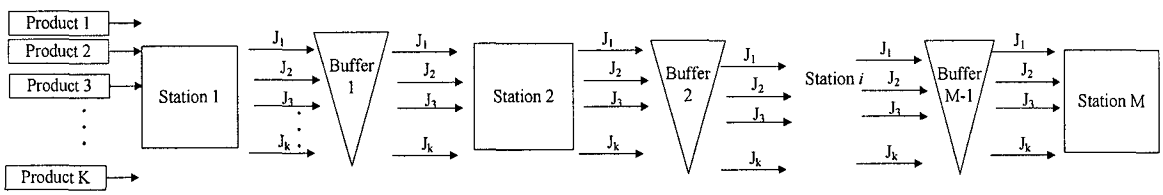

Figure (3.1): A Typical Multi-Product Production Line

Based on the reviewed literature, there is no feasible exact solution for the unreliable production lines with more than two workstations in series and one buffer in the middle; however, near-optimal solution is possible. There exist several approximation methods in the literature (such as decomposition method to measure the Batch Production Rate. However, to the best of the author's knowledge there is no published method to approximate this parameter in a multi product unreliable production environment. This fact is a motivation to present an approximation procedure for the production rate in a multi-product production environment that ultimately will enable us to obtain the main performance measure in this thesis which is Total Cycle Time.

Parameters and variables that are involved in the presented model can be defined as Processing times of each product type at each workstation, Batch sizes (volume) of each product type, Mean Time to Failure and Repair of each workstation, buffer size, variation of processing time, and variations of mean time to failure and repair. The distinction between parameters and variables in such models is not clear cut, and it depends on the context in which the variable appears (Bard, 1974). During numerical experiments that will be presented in the next chapter, some of parameters and variables can interchangeably be used depending on target of study and expected conclusions.

In continue, definition of terms, notations, and assumptions are presented and then a solution approach to the described problem will be proposed.

3.3 Definition of terms:

Processing Time:

Processing time is the length of time a part resides at a workstation to be processed. The Processing Rate is the reciprocal of the processing time.

Total Cycle Time:

Total Cycle Time is the summation of all Batch Cycle Times in the system. That is the time required for producing all batches of products that are scheduled for processing.

Line set up time:

That is the time that a batch spends waiting for the whole line to be set up. It may include change of tools, fixtures, adjustments and coordination of machines and material handling systems.

Part Throughput time:

Part throughput time is the time that each part spends in the system (from entering the first workstation to exiting the last workstation.

Production Rate:

This measure of performance is the rate at which parts exit from the last workstation of the flow line.

Inter-departure time:

That is the time between two consequent parts exiting from the last workstation.

Blocking of workstation:

Blocking of a workstation is a state that a machine has no space to discharge a completed part. The blocked machine is stopped from processing the next part until a space becomes available in the downstream buffer.

Starvation of machine:

A machine is said to be starved if the machine has no part to process. This arises when the upstream machine is empty. The starved machine is prevented from processing until a new part arrives in the upstream buffer.

Flow line:

In this thesis, the term flow line refers to serial production line (Askin & Standridge, 1993). Also, for the sake of simplicity the terms flow line and production line are interchangeably used in this thesis.

Performance measure:

Measures of performance in a manufacturing system are metrics that tell managers what actually has happened in the system in order to control and make appropriate decisions to improve the performance of the system. Among the most important measures of performance, production rate and average number of parts stored in buffers can be named (Dallery and Gershwin 1992).

Queuing Models:

In a simple but typical queuing model, customers arrive from time to time and join a queue (waiting line), are eventually served and finally leave the system. The term "customer" refers to any type of entity that can be viewed as requesting "service" from a system (Banks et al. 2005). In the presented model at this thesis, each workstation can be viewed as server and each part of any product type as customers. Typical measures of system performance include server utilization (percentage of time server is busy), length of waiting line, and delays of customers.

Queuing Notation:

Recognizing the diversity of queuing systems, Kendal (1953) proposed a notational system for parallel server systems which has widely been adopted. An abridged version of this convention is based on the format A/B/c/N/K. These letters represent the following system characteristics:

A represents the interarrival-time distribution.

B represents the service-time distribution.

c represents the number of parallel servers.

N represents the system capacity (the maximum number of customers in the system). In this thesis, letters N and Z are used interchangeably.

K represents the size of the calling population (the population of potential customers).

Common symbols for A and B include M (exponential or Markov), D (constant or Deterministic), E_k (Erlang of order k), PH (phase-type), G (arbitrary or general), and GI (general independent).

Fore example, M/M/1/ indicates a single-server system that has unlimited queue capacity and an infinite population of potential arrivals. The interarrival times and service times are exponentially distributed. A GI/G/1/N queue which will be used in this thesis indicates a system with interarrival times that are generally (arbitrarily) and independently distributed and service times that are generally distributed. One server in parallel exists in the system and the total number of customers that can be waiting in the line is N.

3.4 Notation

In this section, we present the notations we frequently use in this thesis.

Subscript i denotes workstations in series.

Where $i = 1, 2, 3, \ldots\ldots, M$

Subscript k denotes product type

Where $k = 1, 2, 3, \ldots\ldots K$. (K is the total number of products)

J_k: Batch size (volume) of product type k

Where $J_k = J_1, J_2, J_3 \ldots J_K$

Subscript n_k denotes the n_{th} part of batch k

Where $n_k = 1, 2, 3\ldots J_k$

T_{nk}: Exit time of the n_{th} part of batch k from the system

Subscript n denotes the n_{th} part of each batch

b_{ki}: Processing time of each part of batch k at work station i

b_i: Mean processing time at workstation i

$CV^2{}_i$: Squared coefficient of variation of the processing time of workstation i

$b_{c,\ i}$: Mean completion time of a part at workstation i after taking system failure and repair into consideration

$CV^2{}_{c,\ i}$: Squared coefficient of variation of the processing time at workstation i after taking system failure and repair into consideration

R_{ki}: Processing rate of product type k at work station i

μ_i: Mean processing rate at station i

THR_{nki}: Throughput time of the n_{th} part of batch k at work station i

THR_{nk}: Throughput time of the n_{th} part of batch k at the entire line

TC: Total cycle time. That is the duration of time from entering of the first part of the first batch into the system to the time that the last part of the last batch leaves the system.

BP_k: Processing time of batch k. That is time lag from the start of the process of the first part of batch *k* at the first workstation to the finish of the process of the last part at the last workstation.

BC_k: Cycle time of batch k. That is Batch Processing Time plus set up time required for the line before it starts to process a batch.

SU_k: Time required for the set up of the production line to process product type *k*.

T_k_out: Clock time, when the last part of batch k exits the last machine

T_k_in: Clock time, when the first part of the batch enters the first machine in the line

ID_k: Inter-departure time between each part(pieces) of batch k

X_k: Production rate of batch k

X: General Production Rate

S_i: Workstation *i*

M: Total number of workstations composing the production line;

$B_{i, i+1}$: Buffer between stations *i* and *i+1*;

$S_u(i, i+1)$: Pseudo upstream station; that is the segment of the production line upstream of buffer $B_{i, i+1}$;

$S_d(i, i+1)$: Pseudo downstream station; that is the segment of the production line downstream of buffer $B_{i, i+1}$

L(i, i+1): Subsystem of the production line L(i, i+1) is made of two pseudo stations $S_u(i, i+1)$ and $S_d(i, i+1)$ as well as a pseudo buffer B(i, i+1);

S_u (i, j): Pseudo upstream station in the subsystem L (i, i+1)

S_d (i, j): Pseudo downstream station in the subsystem L (i, i+1)

μ_u (i, j): Processing rate of pseudo upstream station S_u (i, j);

μ_d (i, j): Processing rate of pseudo downstream station S_d (i, j);

P_b (i, j): Blocking probability of pseudo upstream station S_u (i, j);

P_s (i, j): Starving probability of pseudo downstream station S_d (i, j);

CV_u (i, j): Coefficient of variation of the processing time of pseudo upstream station S_u (i, j);

CV_d (i, j): Coefficient of variation of the processing time of pseudo downstream station S_d (i, j);

$\rho_{(i,i+1)}$: Utilization factor of Subsystem L(i, i+1)

λ_i: Failure rate of workstation *i*

$MTTR_i$: Mean time to repair of workstation *i,* S_i;

$CV^2_{R,i}$: Squared coefficient of variation of the time to repair of station S_i;

$MTTF_i$: Mean to failure of workstation *i,* S_i;

$C_{i,\, i+1}$: Size of buffer $B_{i,\, i+1}$

$Z_{i,\, i+1}$: Maximum capacity of subsystem L (i, i+1); that is ($C_{i,\, i+1}$ + 1)

$\widehat{N_{(i,i+1)}}$: Mean number of parts in the subsystem L(i, i+1).

X_{old}: Production rate of the preceding iteration;

X_{new}: Production rate of the actual iteration;

ϵ: Convergence tolerance.

3.5 Assumptions

- First machine never starves, so there are always parts to produce.

- Automated material handling (such as conveyors) is used to transfer parts between machines. They are assumed to be reliable.
- Products pass through all of the stages of the line as indicated by the arrows in Figure 3.1.
- Backward in the flow line is not allowed.
- Parts arrivals are in batches, but they enter the production line one by one.
- Batches are served on a First Come First Served discipline, but within each batch, parts are chosen on a random base.
- Each workstation has an exponential Mean Time to Failure (MTTF) and a generally distributed Mean Time to Repair (MTTR).
- When a machine failure occurs, the part that is being processed remains at the same machine during repair and the remaining operations are done when the machine is up. So, there is no rework or scrap when a machine breaks down.
- Buffers with finite sizes are located between workstations.
- When the processing of one product type is finished, the whole flow line requires set up before the production of a new product is started.
- Line set up time is deterministic
- Transportation time of parts between workstations is included in the processing time.

3.6 Solution methodology

As mentioned earlier, an approximation method is proposed to evaluate the Total Cycle Time of the flow line as a performance measure. A brief description of this approach is given at first, and an in-depth and more detailed explanation comes after.

3.6.1 Summary of the solution approach

As per definition, Total Cycle Time is a summation of all Batch Cycle Times in the system:

$$TC = \sum_{k=1}^{K} BC_k \tag{1}$$

And Batch Cycle Time is obtained by:

Batch Cycle Time (of batch k) = Set up Time (of batch k) + Batch Processing Time:

$$BC_k = SU_k + BP_k \tag{2}$$

Batch Processing Time (BP_k) is the time lag from the start of processing the first part at the first workstation until the completion of the last part at the last station. That is:

$$BP_k = T_{k}\_out - T_{k}\_in \tag{3}$$

We can obtain this parameter as explained below. Figure (3.2) gives a better understanding of the Batch Processing time. T_{1k} is the time that first part of batch k exits the line, and T_{2k} is the exit time of second part, and so on. Processing of the batch starts from time 0 and ends at time T_{Jk} (T_{Jk} is the time that last part of batch k exits the line).

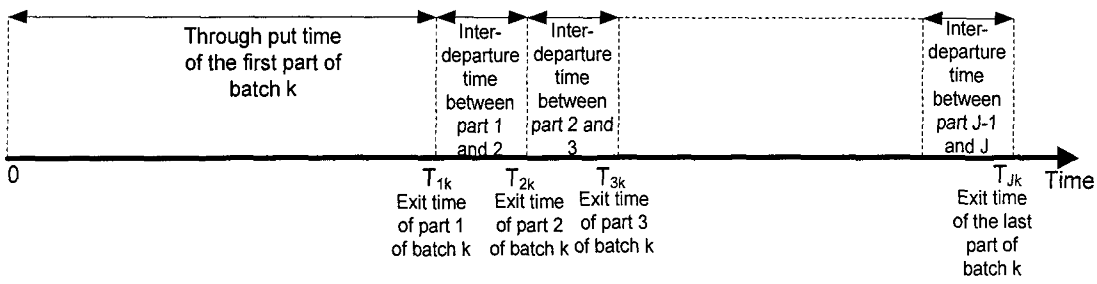

Figure (3.2): Analysis of Batch Processing Time

Therefore, the Batch Processing Time, consists of:

Throughput time of the first part + time between departure of the first and second part + time between departure of the second and third part + … + time between departure of the J_{th} and $(J-1)_{th}$ part. Or:

Batch Processing Time (BP_k): $T_{Jk} = T_{1k} + (T_{2k} - T_{1k}) + (T_{3k} - T_{2k}) + \ldots + (T_{Jk} - T_{Jk-1})$ (4)

Batch Processing Time = Throughput time of the first part of the batch + (J_k-1) * (inter-departure time), or:

$$BP_k = THR_{1k} + (J_k - 1) * ID_k \qquad (5)$$

The time that a part spends in the system is the summation of times that it spends in each workstation (based on assumption, transportation time is included in part processing time). In other words, throughput time of each part is the summation of throughput time of the part in each workstation:

$$THR_{1k} = THR_{1k1} + THR_{1k2} + THR_{1k3} + \ldots + THR_{1km}$$

Therefore, we can rewrite equation (5) as:

$$BP_k = \sum_{i=1}^{m} THR_{1ki} + (J_k - 1) * ID_k \qquad (6)$$

We also know that:

Inter-departure time = 1 / Batch Production Rate

Thus:

Batch Processing Time = Throughput time of the first part of the batch + (J-1) / (Batch Production Rate), or:

$$BP_k = \sum_{i=1}^{m} THR_{1ki} + (J_k - 1) / X_k \qquad (7)$$

Due to the high variety of product types in the real world, it may not be practical or even feasible to calculate the parameters THR_{1ki} and X_k for every product type. Also, performance measure X_k should not change when the product type changes; it should stay unique throughout the production of all product types.

To solve this issue, we propose Generalization of the two parameters THR_{1ki} and X_k by introducing the notion of General production rate (X) instead of X_k, and General part throughput time (THR_{1i}) instead of THR_{1ki}. This way, we do not need to measure batch production rate X_k for every product type (it will be independent of product type). Instead, we have a general production rate for the whole production system. This is important especially when the product variety is high, production volume is low, and there is uncertainty in production of some product types during each period.

Therefore, equation (7) can be rewritten as:

$$BP_k = \sum_{i=1}^{m} THR_{1i} + (J_k-1)/X \tag{8}$$

Where, THR_{1i} and X are general measures needed to calculate BP_k and J_k is the batch size (volume) of product type k.

By substituting the above equation in equations (2) and (1), we will have:

$$BC_k = SU_k + \sum_{i=1}^{m} THR_{1i} + (J_k - 1) / X \tag{9}$$

Thus:

$$TC = \sum_{k=1}^{K} SU_K + \sum_{k=1}^{K} \sum_{i=1}^{m} THR_{1i} + \sum_{K=1}^{k} ((J_k - 1)/X) \text{ Or:}$$

$$TC = \sum_{k=1}^{K} SU_K + K * \sum_{i=1}^{m} THR_{1i} + (1/X) * \sum_{K=1}^{k} (J_k - 1) \tag{10}$$

We will explain how to calculate parameter X later in this chapter. It will be shown that for the calculation of batch production rate X, several input parameters are needed, including part processing time b_{ki}. For now, we focus on how to generalize the parameters X_k and THR_{1ki}.

In order to answer the above question, first we should answer another question: What are the sources of variability when product type changes in the system, or simply, what changes happen in the system if a product type changes? To answer this question we should say "processing time at each workstation", and the "flow line set up time" are two changes that are directly related to the product type changes. if product type changes. We disregard line set up time for now because it doesn't affect the calculation of X and THR_{1i}. Therefore, we can state that a change in product type is directly reflected by a change in the part processing time. In another word, part processing time in each work station is the source of variability related to the product type changes, and if we tackle this then in fact, we are able to tackle the variety in product type. As a result, we will be able to find a general value for X and THR_{1i}.

Now, another question arises, and that is how to mathematically cope with the variability in part processing time at each workstation? In order to answer this question, we propose to approximate this variability by a random variable that is generally (arbitrarily) distributed. Parameters of this distribution are mean (b_i) and coefficient of variation (CV_i), where b_i is the mean of processing time of all product types at workstation *i*, and parameter CV_i is coefficient of variation of part processing time of all product types at workstation *i*.

Now that we are able to tackle the source of variability related to product type changes, by using generalized "part processing time" with parameters b_i and CV_i, we will be able to obtain general values for parameters X and THR_{1i}. The following sections of this chapter will describe this procedure in detail.

3.6.2 Detailed description of the solution approach

3.6.2.1 Computing Throughput Time of the first part at workstation *i* (THR_{1ki})

If we assume buffer as a finite queue (figure 3.3) and each workstation as a server, the throughput time is the summation of "time spent in the queue" and "time spent at workstation":

THR_{1ki} = Time spent in the queue (buffer) + Time spent at workstation

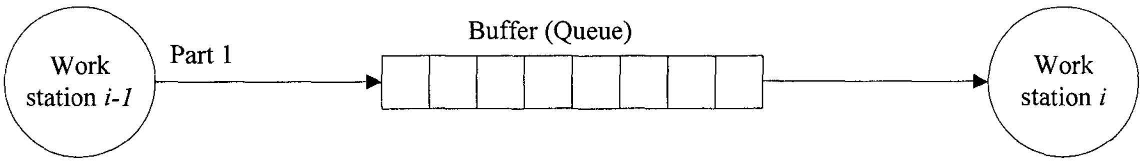

Figure (3.3): Demonstration of the queue between two subsequent workstations

Now, because we are studying the first part of each batch that arrives into the system, the "Time spent in the queue" is equal to zero. Therefore:

THR_{1ki} = Time that the first part of batch k spends at workstation *i*

For the calculation of "Time spent at workstation" we should take the unreliability of the workstation into consideration. Based on assumptions, when a machine failure occurs, the part that is being processed remains at the same machine during repair and the remaining operations are done when the machine is up. Thus:

Time spent at workstation i =

Processing time of the part at workstation i +

Number of failures during the processing time * Mean Time to Repair

If λ_i is the failure rate of workstation i, and b_{ki} represents the processing time of each part of batch k at workstation i, then we have:

Time spent by each part of batch k at workstation $i = b_{ki} + \lambda_i * b_{ki} * MTTR_i$

And since based on assumption again, the failure of workstations follows an exponential distribution, λ_i equals to reciprocal of Mean Time to Failure ($MTTF_i$), and we can rewrite the above equation as:

$$THR_{1ki} = b_{ki} + b_{ki} * (MTTR_i/MTTF_i) \tag{11}$$

As explained earlier, if we assume that part processing time has a general distribution with mean of b_i at each workstation, we can obtain general THR_{1i} by replacing b_{ki} with b_i as below:

$$THR_{1i} = b_i + b_i * (MTTR_i/MTTF_i) \tag{12}$$

3.6.2.2 Computing General Production Rate (X)

In the next step, we need to find a solution for measuring the general batch production rate (X). For this purpose, we adapt an algorithm based on the decomposition method of Tempelmeier and Burger (2001).

At the same time, we will apply GI/G/1/N queuing models as per Buzacott and Shanthikumar (1993) and making the necessary modifications to be able to calculate parameters (such as blocking and starving probabilities) that are needed for the above-mentioned algorithm. The following subsection presents the detailed description of the above-mentioned procedure to calculate the Aggregate Production Rate.

3.6.2.2.1 Model description

We consider a production line including M single-server stations separated by M-1 finite buffers. Discrete parts enter the system at station 1 and, consecutively proceed from station 1 to station M in a fixed predetermined sequence. Storage space in front of the first station is unlimited and never empty (i.e. station 1 is never starved), and the last station has always enough space to unload a completed part (i.e., station M is never blocked). The system has the following characteristics:

- Random processing times: the processing times at station *i* are generally distributed random variables with mean b_i and coefficient of variation CV_i.
- Station breakdowns: Failures can occur at any point in time when a station S_i is operating, i.e., when it is never starved or blocked. The repair time is characterized with its mean $MTTR_i$ and its coefficient of variation $CV_{R,i}$. The time to failure is assumed to be exponentially distributed with mean $MTTF_i$.

The flow line is modeled as a system of finite queues as shown below:

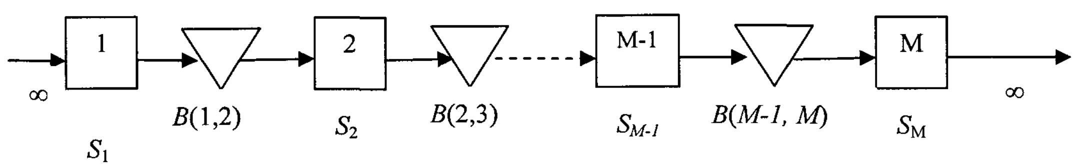

Figure (3.4): A production line including M single-server stations M-1 finite buffers

The buffer $B_{i,\, i+1}$ located between stations S_i and S_{i+1} has a capacity of $C_{i,i+1}$ parts.

3.6.2.2.2 Decomposition procedure of reliable production line

At first and as a building block, a reliable production line is considered and later on, station failure will be taken into consideration. The flow line consisting of M stations is decomposed into M-1 two-station subsystems. Each subsystem L(i, i+1) is made-up of two pseudo stations $S_u(i, i+1)$ and $S_d(i, i+1)$ as well as a pseudo buffer B(i, i+1).

The pseudo station $S_u(i, i+1)$ represents the segment of the flow line upstream of buffer $B_{i,\, i+1}$, whereas the pseudo station $S_d(i, i+1)$ represents the segment of the flow line downstream of buffer $B_{i,\, i+1}$. Pseudo station $S_u(i, i+1)$ is never starved, but blocked with the probability $P_b(i, i+1)$. Pseudo station $S_d(i, i+1)$ is never blocked, but starved with the probability $P_s(i, i+1)$.

Parameters $\mu_u(i, i+1)$ and $\mu_d(i, i+1)$ are the processing rates of upstream and downstream of sub-system L(i, i+1), which include the effects of starving of all upstream stations and the effects of blocking of all downstream stations. The decomposed system is shown below in Figure 3.4 and 3.5:

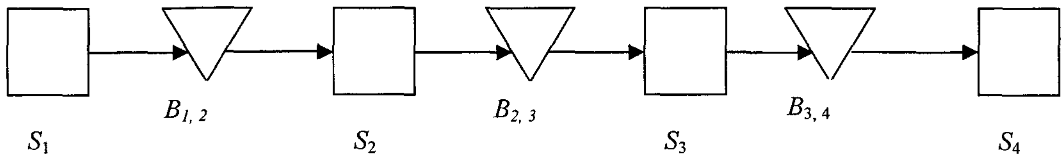

Figure (3.5): Original system

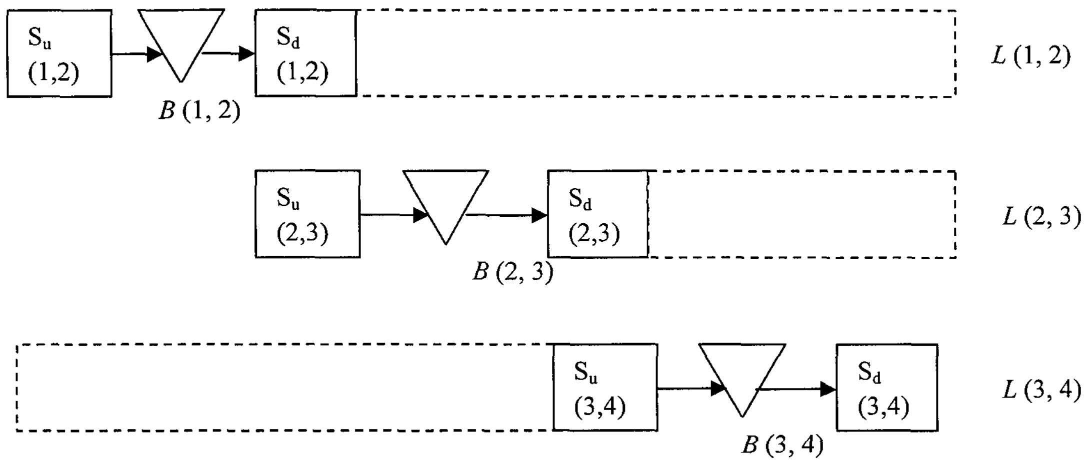

Figure (3.6): Decomposed system

The adjusted mean processing rate of the downstream and upstream stations (including blocking and starving) is given by equations (1), and (2).

$$\mu_d(i, i+1) = \mu_{i+1} * (1-P_b(i+1, i+2)) \tag{13}$$

$$\mu_u(i, i+1) = \mu_i * (1-Ps(i-1, i)) \tag{14}$$

The impact of blocking and starving on the coefficient of variation of the adjusted processing times is modeled using the approximation proposed by Buzacott et al. (1995) as below:

$$CV_d^2\,(i-1,i) = \mu_d^2(i-1,i)\left(\frac{CV_i^2+1}{\mu_i^2} + \frac{P_b(i,i+1)}{X\,(i,i+1)} * \left[\frac{2}{\mu_i} + \frac{CV_d^2(i,i+1)+1}{\mu_d(i,i+1)}\right]\right) - 1. \tag{15}$$

$$CV_u^2\,(i,i+1) = \mu_u^2(i,i+1)\left(\frac{CV_i^2+1}{\mu_i^2} + \frac{P_s(i-1,i)}{X\,(i-1,i)} * \left[\frac{2}{\mu_i} + \frac{CV_u^2(i-1,i)+1}{\mu_u(i-1,i)}\right]\right) - 1. \quad (16)$$

The production rate X (i, i+1) of subsystem L(i, i+1) is equal to the processing rate of station Su(i, i+1), if Su(i, i+1) is not blocked and it is equal to the processing rate of station Sd(i, i+1) if Sd(i, i+1) is not starved. Therefore, we have:

$$X\,(i, i+1) = \mu_u\,(i, i+1) * [1 - P_b\,(i, i+1)]$$

$$= \mu_d\,(i, i+1) * [1 - Ps\,(i, i+1)] \quad (17)$$

Moreover, it is equal to the processing rate of station S_i, if it is in operating condition. This is true, if S_i is neither blocked nor starved.

$$X\,(i, i+1) = \mu_i * [1 - P_b\,(i, i+1)] * [1 - Ps\,(i-1, i)]$$

$$= \mu_{i+1} * [1 - P_b\,(i+1, i+2)] * [1 - Ps\,(i, i+1)] \quad (18)$$

The blocking and starving probabilities are determined based on the approximation for a $GI/G/1/Z_{max}$ queuing system given in Buzacott and Shanthikumar (1993).

The probabilities of starving and blocking of the stations in the subsystem L(i, i+1) are presented below:

$$P_s\,(i, i+1) = \frac{(1-\rho_{(i,i+1)})}{(1-\rho_{(i,i+1)}*\sigma_{(i,i+1)}{}^{z_{i,i+1}})} \quad (19)$$

$$P_b\,(i, i+1) = \frac{\rho_{(i,i+1)}(1-\sigma_{(i,i+1)})\sigma_{(i,i+1)}{}^{z_{i,i+1}-1}}{(1-\rho_{(i,i+1)}*\sigma_{(i,i+1)}{}^{z_{i,i+1}})} \tag{20}$$

Where:

$$\rho_{(i,i+1)} = \frac{\mu_u\,(i, i+1)}{\mu_d\,(i, i+1)} = system\ utilization$$

$$\sigma_{(i,i+1)} = \frac{(\widehat{N_{(i,i+1)}} - \rho_{(i,i+1)})}{\widehat{N_{(i,i+1)}}}$$

Note: $\sigma_{(i,i+1)}$ is a dimensionless factor defined for the ease of calculations.

$$\widehat{N_{(i,i+1)}} = \left\{\frac{\rho_{(i,i+1)}{}^2(1 + CV_d^2(i, i+1))}{1 + \rho_{(i,i+1)}{}^2 * CV_d^2(i, i+1)}\right\} * \left\{\frac{CV_u^2(i, i+1) + \rho_{(i,i+1)}{}^2 * CV_d^2(i, i+1)}{2(1 - \rho_{(i,i+1)})}\right\} + \rho_{(i,i+1)}$$

Where $\widehat{N_{(i,i+1)}}$= Mean number of parts in the system L(i, i+1).

3.6.2.2.3 Consideration of station failures

The approximation outlined in the previous subsection can be extended to analyze flow lines with unreliable stations. Operation dependent failures are incorporated with the help of the *completion time concept*. By definition presented in Gaver (1962), completion time is the duration of the period that elapses between the instant at which the service of an Arrival (of the *n*th part) begins and that at which process of the next (n+1)th does begin.

In our model, if a failure occurs at station S_i, the current operation is interrupted and the currently processed part waits at the station until the repair has been finished and then, processing can be continued. Because the processing time and the time to failure are random variables, the station may break down again before it has finished the current part. Theoretically, an infinite number of breakdowns are possible. The length of time the part resides at the station is thus made up of a "processing" (up time) and a "repair"' (down time).

The mean completion time of a part at station (i) is then given by the following Equation:

$$b_{C,i} = b_i + \frac{MTTR_i}{MTTF_i} * b_i \tag{21}$$

The coefficient of variation of the completion time depends on the distributions of the times-to-failure and the repair times. Assuming an exponential distribution for the times-to-failure and generally distributed repair times with mean $MTTR_i$ and coefficient of variation $CV_{R,i}$ then we get, according to Tempelmeier and Burger (2001):

$$CV_{C,i}^2 = CV_i^2 + \frac{MTTF_i * (CV_{R,i}^2 + 1)}{b_i} \Big/ \left(1 + \frac{MTTF_i}{MTTR_i}\right)^2 \tag{22}$$

3.6.2.2.4 Algorithm (1): General Production Rate calculation

The following unknown parameters μ_d (i, i+1), μ_u (i, i+1), CV^2_u (i, i+1), CV^2_d (i, i+1), P_s (i, i+1), P_b (i, i+1) are computed iteratively with the help of the aforementioned equations. After initialization in step_1, the blocking and starving probabilities are updated in alternating forward and backward passes in step_2.

The values of P_s (i, i+1) are actualized in a forward pass using the values of μ_u (i, i+1) and μ_d (i, i+1), and subsequently in a backward pass the P_b (i, i+1) values are updated. In the following, the unknowns are recalculated using the current values of all parameters updated in the preceding calculations. The approximation terminates if changes of the production rates are within a given tolerance ϵ. The algorithm is summarized below.

```
Procedure main
        Begin
                Call step_1
                If M=2 then
                        Use GI/G/1/Z approximation to compute Ps (1, 2)
                        Xnew = μ2 * (1- Ps (1, 2))
                else
                        Do while abs (Xnew - Xold) < ϵ
                                Xold = Xnew
                                Call step_2
                                Xnew = μM * (1 - Ps (M-1, M))
                        Loop
                End if
                Output Xnew
        End

        Procedure step_1        (*Initialization*)
        Begin
                Input parameters:
                                {M, μi, MTTFi, MTTRi, CV²R,i, Zi,i+1  ∀ i = 1, … , M-1}
                Xold = 0
```

$X_{new} = \mu_M$

$\epsilon = 1/10000$

For i = 0 to M

Ps (i, i+1) = 0

Pb (i, i+1) = 0

$$CV^2_i = CV^2_i + \frac{MTTF_i * \mu_i * (CV^2_{R,i} + 1)}{(1+(\frac{MTTF_i}{MTTR_i}))^2}$$

$$\mu_i = \frac{\mu_i}{(1+\frac{MTTR_i}{MTTF_i})}$$

Next i

End

Procedure step_2 (*iteration*)

Begin

For i = 1 to M-1 (*forward pass*)

μ_u (i, i+1) = μ_i * [1 - P_s (i-1, i)]

μ_d (i, i+1) = μ_{i+1} * [1 – P_b (i+1, i+2)]

if i > 1 then

Compute CV^2_u (i, i+1) with equation (4)

End if

Use GI/G/1/Z approximation to compute P_s (i, i+1)

Next i

For i = M-1 to 1 (*backward pass*)

μ_u (i, i+1) = μ_i * [1 - P_s (i-1, i)]

μ_d (i, i+1) = μ_{i+1} * [1 – P_b (i+1, i+2)]

if i > 1 then

Compute CV^2_u (i, i+1) with equation (4)

End if

Use GI/G/1/N approximation to compute P_b (i, i+1)

Next i

End

3.6.2.3 Calculation of Total Cycle Time

Now that the generalized batch production rate (X) has been obtained, we can continue to calculate the Total Cycle Time, which is the target of the proposed methodology, by using equations (1), (2), and (8). In the next chapter, an extensive numerical study of the applicability, accuracy, and limitations of the proposed method is presented.

CHAPTER 4

NUMERICAL RESULTS

4.1 Introduction

In chapter 3, we presented a solution approach to approximate the performance of a multi product production line as a performance measure. In this chapter, a numerical study will be performed to evaluate the accuracy of the proposed method under different scenarios, analyzing the results, and to draw the appropriate conclusions about the proposed method.

At first and in order to test the accuracy of the proposed method, we will perform a comparison study. As mentioned before, in the literature, we haven't came across to an analytical method for solving a problem similar to what is described in this thesis. Therefore, we need to use simulation as a comparison tool and to evaluate our approximation method.

We examine the proposed method in a variety of cases: at the beginning and for examining the accuracy of algorithm (1) to approximate general production rate X (explained in chapter 3), we consider a single product production line with general processing time at each workstation (case 1). Then we start examining multi-product production lines. We have chosen a variety of different cases, from small production lines to longer production lines and from 2 product types to 5 product types with either deterministic or generally distributed processing times.

Buffer size is a very influential factor in the performance of a production line; In order to study the effect of buffer size on the accuracy of our method, we have performed our experiments based on a wide range of different buffer sizes. The range in most, not all, of the cases is between buffer sizes 1 to 50.

The next step in this chapter will be analyzing the numerical results, and to discuss the strengths and limitations of our method.

As a part of the approximation method, and for the ease of calculations, we have coded the algorithm that is developed to calculate general production rate X (described in chapter 3) in MATLAB. After obtaining X, we will be able to calculate TC using equations (8), (2), and (1) of chapter 3, and then compare them with the simulation results.

We use ProModel software for the simulation purpose in this thesis. The definition of Total Cycle Time in simulation models is same as it was defined in Chapter 3: the duration of time from entry of the first part of the first batch into the system to the time that the last part of the last batch leaves the system. Unit of time used in these experiments is, unless otherwise specified, Time unit (of the simulation software).

Percentage difference is calculated as follow all over this thesis:

$$\%\text{Difference} = \frac{\text{Simulation} - \text{Thesis Approximation}}{\text{Simulation}} \times 100$$

Mean Time to Failure and Repair (MTTF and MTTR) throughout this study is assumed to be exponential.

Quantities of interest are general Batch Production Rate (X), general Throughput time at each workstation (THR_i), general Batch Processing time (BP), and Total Cycle time (TC). The results of approximation method, simulation, and the comparison are given below.

Note: During the numerical study and for computation of parameters "Mean processing time (b_i)" and "Coefficient of Variation of processing time (CV_i)", the "sample mean" and "sample variance" equations that are very common in statistics are used. As a reminder and for the ease of readers, the typical equations are shown below:

$$\overline{X} = \frac{\sum_{j=1}^{k} f_j * X_j}{n}$$

$$S^2 = \frac{\sum_{j=1}^{k} f_j * X_j - n\overline{X}^2}{n-1}$$

Where k is the number of distinct values of X and f_j is the frequency of value of X (parameters used in the above equations are typical and not as a part of notation of this thesis).

4.2 Numerical experiment:

Case 1: Validating the accuracy of algorithm (1)

The goal of this study is to study the accuracy of the parameter X (general production rate) obtained by algorithm (1) as explained in chapter 3.

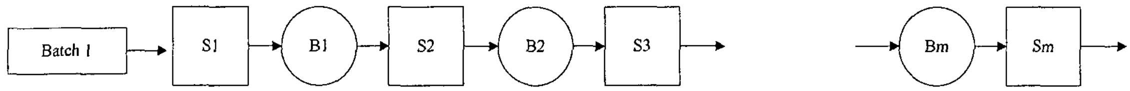

Figure 4.1: Single product 5-workstation production line

A production line as shown in figure (4.1) consisting of 5 identical workstations in series has been considered. For the ease of calculations at the beginning, only one product type has been chosen to omit the effect of set up time. Also, the processing time at each workstation is identical and generally distributed.

In table 4.1 frequency of different values of processing times in a selected workstation (workstations are identical) are shown.

Processing time	Frequency (Percentage)
0.6	8
1.2	52
1.8	8
2	8
2.4	4
3.6	8
1.2	6
1.8	2
2	4

Table 4.1: Distribution of processing time at each workstation

Parameters used in this case:

Time Unit = second

Mean Processing Time b_i (identical for all machines) = 1.548

Mean Processing Rate (μ_i) = 0.6459

Coefficient of Variation of processing time (CV_i) = 0.476611

Squared Coefficient of Variation of processing time ($CV2_i$) = 0.2272

Squared Coefficient of Variation of repair time (CV2r) = 1

Summary of parameters:

	Machine 1	Machine 2	Machine 3	Machine 4	Machine 5
μ_i	0.6459	0.6459	0.6459	0.6459	0.6459
CV2	0.2272	0.2272	0.2272	0.2272	0.2272
MTTF	100	20.1	100	33.34	100
MTTR	10	2.665	10	3.34	10
CV2r	1	1	1	1	1

Table 4.2: Summary of the system parameters for case 1

The parameter X resulted from simulation, thesis approximation, and % difference are presented in table (4.3).

The buffer size at the beginning is considered 1. After buffer size 2, and up to buffer size 10, it increases by increments of 2. The reason is that the effect of any changes in low buffer sizes is significant in the performance of the line. After buffer size 10, the increment will be 4, and it continues up to 42.

This case is simulated for a period of 130,000 hours with a transient period of 30,000 hours for one replication. No statistics are collected during the transient period to ensure the steady-stat behavior of the system.

Buffer size	Parameter X (Proposed approximation)	Parameter X (Simulation)	%difference = ((Simulation-Approximation)/Simulation)*100
1	0.2464	0.4012	38.58
2	0.3285	0.4327	24.08
4	0.4118	0.4676	11.94
6	0.4532	0.4881	7.15
8	0.4809	0.5033	4.45
10	0.4970	0.5129	3.10
14	0.5167	0.5281	2.16
18	0.5284	0.5373	1.66
22	0.5360	0.5445	1.56
26	0.5414	0.5496	1.50
30	0.5453	0.5537	1.53
34	0.5483	0.5564	1.46
38	0.5506	0.5590	1.51
42	0.5524	0.5613	1.58

Table 4.3: General production rate (X) for case study 1 (single-product 5-workstation production line)

Case 2: 2-product 5-machine production line with deterministic processing time

In this study we compare Total Cycle Time obtained by the approximation method presented in chapter 3 with simulation results. The studied production line consists of 5 machines and 2 product types that have deterministic part processing time b_{ki} at each workstation. This case is simulated for a period of 88,400 hours with a transient period of 13,400 hours for one replication.

Processing times at each workstation, line setup time, and batch sizes of each product type are given in tables 4.4 and 4.5. Workstations' Mean Time to Failure and Repair are also given in table 4.6.

In this experiment, the buffer size at the beginning is considered 1. After buffer size 2, and up to buffer size 10, it increases by incremental of 2. After buffer size 10, the increment will be 4, and it continues up to 42.

Time Unit = Second

	Product	Workstation				
		1	2	3	4	5
Processing Time	1	80	70	60	75	65
	2	60	90	70	55	85

Table 4.4: Processing times (Time unit) at each workstation

	Product	Batch Size (unit)	Line Set-up time
Lot Size	1	60	1250
	2	75	1250

Table 4.5: Batch sizes and Line set up time (Time unit)

	Workstation				
	1	2	3	4	5
MTTF	1695	1538	1613	1351	1818
MTTR	213	270	182	149	238

Table 4.6: Mean time to failure and mean time to repair for case 2 (Per time unit)

Parameters of the system which are used for the approximation are as follows:

	Workstation				
	1	2	3	4	5
Mean processing time (b_i)	68.8889	81.1111	65.5556	63.8889	76.1111
Mean Processing Rate (μ_i)	0.0145	0.0123	0.0153	0.0157	0.0131
Coefficient of Variation of processing time (CV_i)	0.1448	0.1230	0.0761	0.1561	0.1311
Squared Coefficient of Variation ($CV2_i$)	0.0210	0.0151	0.0058	0.0244	0.0172
Squared Coefficient of Variation of repair time (CV2r)	1	1	1	1	1
MTTF	1695	1538	1613	1351	1818
MTTR	213	270	182	149	238

Table 4.7: Summary of the system parameters for case 2

The results of the approximation method and simulation are presented in table 4.8.

Buffer Size	Thesis, Approximation	Simulation	% difference
1	24689	18516	-33.3387
2	20056	18018	-11.3084
4	18017	17664	-2.0013
6	17274	17565	1.6567
8	16919	17514	3.3951
10	16713	17509	4.5468
14	16489	17514	5.8498
18	16359	17514	6.5931
22	16277	17511	7.0486
26	16219	17498	7.3073
30	16175	17507	7.6089
34	16141	17503	7.7815
38	16114	17505	7.9463
42	16092	17502	8.0562

Table 4.8: Total Cycle Time (time unit=second) for case study 2: (2-product 5-machine production line with deterministic processing time)

Case 3: 2-product 5-workstation production line with identical generally distributed processing time at each workstation:

In this case, we consider a 5-machine production line that produces 2 product types. Each product type has identical generally distributed processing time at each workstation. The processing time distribution of each workstation is presented in table 4.9, and the summary of the remaining parameters are given in table 4.10.

Processing time	Frequency (Percentage)
0.6	8
1.2	52
1.8	8
2	8
2.4	4
3.6	8
1.2	6
1.8	2
2	4

Table 4.9: Distribution of processing time (Time unit) at each workstation

Parameters used in case 3:

Time Unit = second

Mean Processing Time (identical for all machines) = 1.548

Mean Processing Rate = 0.6459

Coefficient of Variation of processing time (CV_i) = 0.476611

Squared Coefficient of Variation of processing time ($CV2_i$) = 0.2272

Squared Coefficient of Variation of repair time (CV2r) = 1

	Machine 1	Machine 2	Machine 3	Machine 4	Machine 5
μ	0.6459	0.6459	0.6459	0.6459	0.6459
CV2	0.2272	0.2272	0.2272	0.2272	0.2272
MTTF	30	90	100	100	100
MTTR	10	8	6	4	2
CV2r	1	1	1	1	1

Table 4.10: Summary of parameters used in case3:

This case is simulated for a period of 90,000 hours with a transient period of 15,000 hours for one replication, and Time unit is considered as second.

The buffer size varies until the results reach almost a plateau.

Table 4.11 below presents the results obtained by the proposed approximation method, and these results are compared to the simulation results for validation purposes.

Buffer size	Thesis, Approximation	Simulation	%difference = ((Simulation-Approximation)/Simulation)*100
1	3104	2861	-8.49353
2	2978	2840	-4.85915
3	2918	2830	-3.10954
4	2886	2825	-2.15929
6	2854	2820	-1.20567
8	2834	2818	-0.56778
10	2822	2818	-0.14194
12	2815	2817	0.070998
14	2809	2817	0.28399
16	2805	2817	0.425985
20	2800	2817	0.603479
24	2797	2817	0.709975
28	2795	2817	0.780973
34	2793	2817	0.85197
44	2792	2817	0.887469
60	2792	2817	0.887469

Table 4.11: Total Cycle Time (time unit=second) for case study 3: 2-product 5-workstation production line with identical generally distributed processing time

Case 4: 2-product 10-workstation production line with identical generally distributed processing time at each workstation:

In this case, we consider a 2-product 10-workstation flow line that each of its workstations has an identical generally distributed service time. The service time distribution of each workstation is presented in the following table:

Value	Percentage
0.6	8
1.2	52
1.8	8
2	8
2.4	4
3.6	8
1.2	6
1.8	2
2	4

Table 4.12: Distribution of processing time at each workstation

Parameters:

Time Unit = minute

Mean Processing Time (identical for all machines) = 1.548

Mean Processing Rate = 0.6459

Coefficient of Variation of processing time (CV_i) = 0.476611

Squared Coefficient of Variation of processing time ($CV2_i$) = 0.2272

Squared Coefficient of Variation of repair time (CV2r) = 1

Summary of machine parameters:

	Machine 1	Machine 2	Machine 3	Machine 4	Machine 5	Machine 6	Machine 7	Machine 8	Machine 9	Machine 10
μ	0.6459	0.6459	0.6459	0.6459	0.6459	0.6459	0.6459	0.6459	0.6459	0.6459
CV2	0.2272	0.2272	0.2272	0.2272	0.2272	0.2272	0.2272	0.2272	0.2272	0.2272
MTTF	30	90	100	100	100	100	100	100	100	100
MTTR	10	8	6	4	2	2	2	2	2	2
CV2r	1	1	1	1	1	1	1	1	1	1

Table 4.13: Summary of parameters used in case 4

This case is simulated for a period of 400,000 hours with a transient period of 100,000 hours for one replication.

As like as case 2, the buffer size at the beginning is considered 1. After buffer size 2, and up to buffer size 10, it increases by increments of 2. After buffer size 10, the increment will be 4, and it continues up to 50 until it reaches a plateau. Time Unit s considered Minute in this case.

Numerical results for Cycle time comparison in the 10-machine production line are presented below:

Buffer size	Thesis, Approximation	Simulation	%difference = ((Simulation-Approximation)/Simulation)*100
1	3076	2892	-6.35
2	2932	2870	-2.15
3	2916	2861	-1.91
4	2895	2857	-1.33
6	2868	2853	-0.52
8	2850	2852	0.07
10	2838	2851	0.47
14	2825	2851	0.91
18	2818	2851	1.17
22	2814	2850	1.28
26	2812	2851	1.36
30	2810	2851	1.43
34	2809	2851	1.47
38	2807	2851	1.53
42	2808	2851	1.50
46	2808	2851	1.50
50	2808	2851	1.50

Table 4.14: Total Cycle Time (time unit=minute) for case study 4: 2-product 10-workstation production line with identical generally distributed processing time

Case 5: 2-product 10 workstation production line with deterministic processing times

In this case, we consider a 10 workstation production line that produces 2 product types; however, each product has deterministic processing time at each workstation. Characteristics of the system including processing time, line set up time, batch size, and Mean Time to Failure and Repair are given in tables 4.15 to 4.17 as below:

Time Unit = Second

	Product	Workstation									
		1	2	3	4	5	6	7	8	9	10
Processing Time	1	80	70	60	75	65	75	70	65	60	60
	2	60	90	70	55	85	55	90	85	70	70

Table 4.15: Processing times (Time unit=second) at each workstation for case 5

	Product	Batch Size (unit)	Line Set-up time
Lot Size	1	60	1250
	2	75	1250

Table 4.16: Batch Sizes and Line set up time (Time unit)

	Workstation									
	1	2	3	4	5	6	7	8	9	10
MTTF	1695	1538	1613	1351	1818	1550	1500	1450	1700	1600
MTTR	213	270	182	149	238	180	170	270	200	160

Table 4.17: Mean time to failure and mean time to repair for case 2 (Per time unit)

Parameters of the system which are used for the approximation are given in table 4.18.

	Workstation									
	1	2	3	4	5	6	7	8	9	10
Mean processing time (b_i)	68.89	81.11	65.56	83.89	76.11	63.89	81.11	76.11	56.56	56.56
Mean Processing Rate (μ_i)	0.0145	0.0123	0.0153	0.0157	0.0131	0.0157	0.0123	0.0131	0.0153	0.0153
Coefficient of Variation of processing time (CV_i)	0.1448	0.123	0.0761	0.1561	0.1311	0.1561	0.123	0.1311	0.0761	0.0761
Squared Coefficient of Variation ($CV2_i$)	0.02097	0.01512	0.00579	0.02438	0.01718	0.02438	0.01512	0.01718	0.00579	0.00579
Squared Coefficient of Variation of repair time (CV2r)	1	1	1	1	1	1	1	1	1	1
MTTF	1695	1538	1613	1351	1818	1550	1500	1450	1700	1600
MTTR	213	270	182	149	238	180	170	270	200	160

Table 4.18: Summary of the system parameters for case 5

This case is simulated for a period of 50,000 hours with a transient period of 10,000 hours for one replication.

The approximation procedure is performed for different buffer sizes until it reaches a near steady state. From 1 up to 10, buffer size increases by increments of 2 and after buffer size 10 the increments will be 4 up to buffer size 58. This procedure continues until a plateau is reached almost at buffer size 126. For the ease of demonstration and because the differences are very small, the results for buffer sizes between 58 and 126 are omitted from table 4.19.

Moreover, simulation is run for the two end sides of the buffer sizes, where changes are influential. For middle buffer sizes, simulation is not necessary as the trend is obvious and reliable based on previous experiments, and results are not influential for later analysis.

Buffer size	Thesis Approximation	Simulation	%difference = ((Simulation-Approximation)/Simulation)*100
1	29280.6	21101	-38.76
2	21657.2	20264.6	-6.87
3	19795.0	19934.2	0.70
4	18954.2	19799.7	4.27
6	18097.3	19659.7	7.95
8	17781.0	19663	9.57
10	17415.6	19653	11.38
14	17216.6		
18	17235.3		
22	17227.2		
26	17176.8		
30	17139.7		
34	17110.8		
38	17087.3		
42	17067.8		
46	17051.3		
50	17037.1		
54	17024.9	19647	13.35
58	17014.3	19651.7	13.42
126	16938.1	19651.7	13.81
130	16936.3	19651.7	13.82

Table 4.19: Total Cycle Time (Time unit=second) for case study 5: 2-product 10-workstation production line with deterministic processing time

Case 6: 2-product 10 workstation production line with deterministic processing times (different MTTF and MTTR)

This case is exactly the same as case 5, but with different Mean Time to Failure (MTTF) and Repair (MTTR). The purpose of this is studying the effect of MTTF and MTTR on the performance and accuracy of the proposed method. Table 4.20 shows parameters of case 6 including new MTTF and MTTR.

	Workstation									
	1	2	3	4	5	6	7	8	9	10
Mean processing time (b_i)	68.89	81.11	65.56	83.89	76.11	63.89	81.11	76.11	56.56	56.56
Mean Processing Rate (μ_i)	0.0145	0.0123	0.0153	0.0157	0.0131	0.0157	0.0123	0.0131	0.0153	0.0153
Coefficient of Variation of processing time (CV_i)	0.1448	0.123	0.0761	0.1561	0.1311	0.1561	0.123	0.1311	0.0761	0.0761
Squared Coefficient of Variation ($CV2_i$)	0.02097	0.01512	0.00579	0.02438	0.01718	0.02438	0.01512	0.01718	0.00579	0.00579
Squared Coefficient of Variation of repair time (CV2r)	1	1	1	1	1	1	1	1	1	1
MTTF	30	90	100	100	100	100	100	100	100	100
MTTR	10	8	6	4	2	2	2	2	2	2

Table 4.20: Summary of the system parameters for case 6

The simulation time however for this case, due to decrease of MTTR, becomes enormously long. Therefore, time unit is changed from second to minute. On the other side, run time is increased to 250,000 hours with transient (warm up) time of 50,000 hours.

Buffer size incremental increase is as in case 5; however, in this case stability of the results is reached at buffer size 18.

Results of approximation method, simulation, and comparison are presented in table 4.21.

Buffer size	Total Cycle time (Approximation)	Total Cycle time (Simulation)	%difference = ((Simulation-Approximation)/Simulation)*100
1	13523	16216	16.61
2	14244	16217	12.17
3	15242	16218	6.02
4	16105	16216	0.69
6	16223	16219	-0.03
8	16227	16215	-0.07
10	16225	16224	0.00
14	16223	16223	0.00
18	16222	16218	-0.03
22	16222	16218	-0.03

Table 4.21: Total Cycle Time (Time unit = minute) for case study 6: 2-product 10-workstation production line with deterministic processing time (different MTTF & MTTR)

Case 7: 5-machine 5-product production line with deterministic processing time

In this case, the studied production line consists of 5 identical machines in series and 5 product types that have deterministic part processing time b_{ki} at each workstation. System characteristics and summary of parameters are presented tables 4.22 to 4.25.

Time Unit = Minute

	Product	Workstation				
		1	2	3	4	5
Processing Time	1	80	70	60	75	65
	2	60	90	70	55	85
	3	70	95	75	65	80
	4	75	80	65	60	75
	5	65	75	60	70	65

Table 4.22: Processing times (Time unit) at each workstation for case 7

	Product	Lot Size (unit)	Line Set-up time
Lot Size	1	60	1250
	2	75	1250
	3	70	1250
	4	90	1250
	5	85	1250

Table 4.23: batch sizes and Line set up time for case 7

	Workstation				
	1	2	3	4	5
MTTF	30	90	100	100	100
MTTR	10	8	6	4	2

Table 4.24: Mean time to failure and mean time to repair for case 7 (Per time unit)

	Workstation				
	1	2	3	4	5
Mean processing time (b_i)	69.6710526	82.0395	65.9211	64.5395	74.0789
Mean Processing Rate (μ_i)	0.01435316	0.01219	0.01517	0.01549	0.0135
Coefficient of Variation of processing time (CV_i)	0.09845799	0.10807	0.08599	0.10617	0.10613
Squared Coefficient of Variation ($CV2_i$)	0.00969398	0.01168	0.00739	0.01127	0.01126
Squared Coefficient of Variation of repair time (CV2r)	1	1	1	1	1
MTTF	30	90	100	100	100
MTTR	10	8	6	4	2

Table 4.25: Summary of the system parameters for case 7

This case is simulated for a period of 400,000 hours with a transient period of 100,000 hours for two replications.

The same incremental increase in buffer sizes is performed in this case up to reaching a near plateau. Results are presented in table 4.26 below.

Buffer size	Total Cycle time (Approximation)	Total Cycle time (Simulation)	%difference = ((Simulation-Approximation)/Simulation)*100
1	37451.85	41558.2	9.880971
2	40604.68	41551.4	2.27843
3	43130.12	41549	-3.80544
4	43568.03	41543.4	-4.87353
6	43359.14	41550.8	-4.35211
8	43148.07	41552	-3.84113
10	43084.99	41555.3	-3.68109
14	43061.83	41561.4	-3.61016
18	43059.89	41548.8	-3.6369
22	43059.9	41546.2	-3.64341
26	43059.89	41541.5	-3.65511
30	43059.89	41541.5	-3.65511
34	43059.89	41541.5	-3.65511

Table 4.26: Total Cycle Time (Time unit = minute) for case study 7: 5-product 5-workstation production line with deterministic processing time

Case 8: 5-machine 5-product production line with deterministic processing time (different MTTF and MTTR)

Almost all of parameters and system characteristics in this case are as same as case 7, but MTTF and MTTR are different. Summary of parameters is presented in table 4.27.

	Workstation				
	1	2	3	4	5
Mean processing time (b_i)	69.6710526	82.0395	65.9211	64.5395	74.0789
Mean Processing Rate (μ_i)	0.01435316	0.01219	0.01517	0.01549	0.0135
Coefficient of Variation of processing time (CV_i)	0.09845799	0.10807	0.08599	0.10617	0.10613
Squared Coefficient of Variation ($CV2_i$)	0.00969398	0.01168	0.00739	0.01127	0.01126
Squared Coefficient of Variation of repair time (CV2r)	1	1	1	1	1
MTTF	1695	1538	1613	1351	1818
MTTR	213	270	182	149	238

Table 4.27: Summary of the system parameters for case 8

Simulation run time for this case is 100,000 hr with warm up period of 10,000 hr. Time unit is considered second for this case. Buffer size incremental increase is as same as previous cases up to buffer size 58, and near plateau is reached at buffer size 162.

Buffer size	Total Cycle time (Approximation)	Total Cycle time (Simulation)	%difference = ((Simulation-Approximation)/Simulation)*100
1	68635	50582	-35.69
2	54650	49078	-11.35
3	50894	48342	-5.28
4	49709	47931	-3.71
6	48167	47575	-1.25
8	47522	47464	-0.12
10	46979	47450	0.99
14	46360	47389	2.17
18	46006	47425	2.99
22	45774	47369	3.37
26	45608	47377	3.73
30	45484	47414	4.07
34	45388	47394	4.23
38	45312	47421	4.45
42	45251	47385	4.50
46	45200	47363	4.57
50	45157	47396	4.72
54	45121	47397	4.80
58	45090	47379	4.83
162	44841	47379	5.36
166	44838	47379	5.36

Table 4.28: Total Cycle Time (Time unit = second) for case study 8: 5-product 5-workstation production line with deterministic processing time

Case 9: 5-machine 5-product production line with deterministic processing time (different batch size)

All the parameters of case 8 remain unchanged for this case. However, in order to examine the effect of batch size on system performance, this case has been tested with different batch sizes. New batch sizes are presented in table 4.29.

	Product	Batch Size (unit)	Line Set-up time
Batch Size	1	240	1250
	2	300	1250
	3	280	1250
	4	360	1250
	5	340	1250

Table 4.29: Batch sizes and Line set up time

For the ease of follow up, parameters of previous case (case 8) are presented again in table 4.30 below:

	Workstation				
	1	2	3	4	5
Mean processing time (b_i)	69.67	82.04	65.92	64.54	74.08
Mean Processing Rate (μ_i)	0.0144	0.0122	0.0152	0.0155	0.0135
Squared Coefficient of Variation ($CV2_i$)	0.00967	0.01166	0.00738	0.01125	0.01124
Squared Coefficient of Variation of repair time (CV2r)	1	1	1	1	1
MTTF	1695	1538	1613	1351	1818
MTTR	213	270	182	149	238

Table 4.30: Summary of the system parameters for case 9

Run time for this case is 50,000 hrs with warm up period of 100,000 hrs and time unit is second. Buffer size increase follows the same procedure of previous cases up to buffer size 58. The steady state is almost reached at buffer size 218.

Buffer size	Total Cycle time (Approximation)	Total Cycle time (Simulation)	%difference = ((Simulation-Approximation)/Simulation)*100
1	252215.2	178818	-41.05
2	195681.7	171114	-14.36
3	180491.9	167234	-7.93
4	175689.9	164426	-6.85
6	169459.1	161866	-4.69
8	166854.4	160351	-4.06
10	164658.1	159988	-2.92
14	162157.2	159422	-1.72
18	160730.4	159642	-0.68
22	159789.8	159117	-0.42
26	159120.1	159208	0.06
30	158619.8	159456	0.52
34	158233	159244	0.63
38	157925.9	159298	0.86
42	157676.7	159183	0.95
46	157470.9	159262	1.12
50	157298.3	159192	1.19
54	157151.7	159102	1.23
58	157025.9	159303	1.43
218	155921.1	159137	2.020865
222	155916.8	159137	2.023565

Table 4.31: Total Cycle Time (Time unit = second) for case study 9: 5-product 5-workstation production line with deterministic processing time

4.3 Analysis of the results:

4.3.1 Effect of increasing product type variety:

Comparing the results generated in case 2 (2-product 5-workstation) with case 8 (5-product 5-workstation) shows that when product type increases, the proposed approximation method demonstrates higher accuracy. Although in very low buffer sizes (lower than 4) due to high instability of the approximation no judgment can be made, for buffer sizes higher than 6, the higher accuracy of case 8 is observed.

The reason for this can be explained as follow: our model is built upon the assumption of processing times being generally distributed. We have used GI/G/1/N queuing model and a decomposition algorithm (Algorithm 1) that assumes generally distributed processing times.

By increasing the product type variety we become closer to the concept of general distribution of processing time. If we have more product types, more arbitrary phases we have; therefore, closer to a General distribution. This can be an explanation for increasing accuracy, when product type increases.

4.3.2 Effect of MTTF and MTTR:

By comparing case 5 with wit 6 (2-product types 10-workstation production line), we observe that accuracy and stability in case 6 is higher than case 5. The only difference in these two cases is the values of MTTF and MTTR.

Before going into further details, here an explanation is needed to study the effect of MTTF and MTTR on the squared coefficient of variation of the completion time explained in chapter 3. The equation is brought to the attention of reader again as below:

$$CV_{C,i}^2 = CV_i^2 + \frac{MTTF_i * \left(CV_{R,i}^2 + 1\right)}{b_i} \Big/ \left(1 + \frac{MTTF_i}{MTTR_i}\right)^2$$

The above equation indicates variability of the Completion Time in each workstation after taking failure of workstations into consideration.

This equation also very well present that MTTF and MTTR have impact on the squared coefficient of variation of completion time, and consequently, on the whole system performance. Here, an investigation of such impact is performed. The following tables demonstrate how the parameters of cases 5 and 6, particularly MTTF and MTTR, affect $CV^2_{c,i}$.

Workstation	Mean processing time (b_i)	Squared Coefficient of Variation ($CV2_i$)	MTTF	MTTR	$CV^2_{c,i}$
1	0.0145	0.02097	1695	213	0.63356
2	0.0123	0.01512	1538	270	0.858886
3	0.0153	0.00579	1613	182	0.513212
4	0.0157	0.02438	1351	149	0.442957
5	0.0131	0.01718	1818	238	0.655448
6	0.0157	0.02438	1550	180	0.551263
7	0.0123	0.01512	1500	170	0.397497
8	0.0131	0.01718	1450	270	0.953318
9	0.0153	0.00579	1700	200	0.582189
10	0.0153	0.00579	1600	160	0.410418

Table 4.32: Calculation of CV2c,i for Case 5 (2-product 10-workstation production line with deterministic processing time)

Workstation	Mean processing time (b_i)	Squared Coefficient of Variation ($CV2_i$)	MTTF	MTTR	$CV^2_{c,i}$
1	0.0145	0.02097	30	10	0.075345
2	0.0123	0.01512	90	8	0.029874
3	0.0153	0.00579	100	6	0.015594
4	0.0157	0.02438	100	4	0.029025
5	0.0131	0.01718	100	2	0.018187
6	0.0157	0.02438	100	2	0.025587
7	0.0123	0.01512	100	2	0.016066
8	0.0131	0.01718	100	2	0.018187
9	0.0153	0.00579	100	2	0.006966
10	0.0153	0.00579	100	2	0.006966

Table 4.33: Calculation of CV2c,i for Case 6 (2-product 10-workstation production line with deterministic processing time with different MTTF and MTTR)

As observed in the above tables (4.32 and 4.33), the parameter $CV^2_{c,i}$ in case 6 is much lower than case 5. As mentioned earlier in chapter 3, the approximation of Buzacott and Shanthikumar (1993) for a GI/G/1/N queuing model has been applied in the proposed method of this thesis. On the other hand, they state that their approximations (of a GI/G/1/N queue) work very well for lower coefficient of variation. When the squared coefficient of variation tempts to be larger, according to Buzacott and Shanthikumar (1993), the approximations can be poor. This can well justify the increase of accuracy of the proposed approximation method in case 6.

In addition, intuitively and statistically speaking, we should expect higher accuracy of approximations when variability of the studied system reduces.

4.3.3 Effect of batch size:

By comparing the results of cases 8 and 9, it is observed that case 9 which has the same parameters of case 8 but with higher batch size, has produced better results. This can be a motivation to claim that larger batch sizes produce better results.

The rationality for this, again, can be related to the queuing concept. All the approximations of the GI/G/1/N queuing model proposed by Buzacott and Shanthikumar (1993) which are used in the thesis proposed method are studied under steady state condition of the system.

Using small batch sizes may result in not reaching a steady state when parts (as customers) arrive into each workstation (as servers). Not reaching the steady state will have the following consequence: The real world queuing model (represented by simulation) does not reach to steady state while our approximations (applied in the thesis method) are based on the steady state of the system. This may cause larger difference between the results of simulation and thesis approximation method as observed in case 8.

4.3.4 Effect of buffer size:

The use of buffer space is an efficient tool that contributes to enhancing the capacity of the production line (Buzacott, 1971). Introducing buffer space decouples successive workstations by allowing them to work independently for a certain period of time. When a workstation in a line fails or has a longer processing time, the workstation upstream of it can still operate until the upstream buffer fills up, and the downstream workstation can still operate until the downstream buffer becomes empty. The use of buffers also helps to reduce the blocking and starvation times of the work centers and thus lowers the cycle time.

In all studied cases during this thesis, increasing buffer size contributed in the accuracy and stability of the results. Especially in lower buffer sizes, adding extra capacity has great impact on improving the results. In case 8 for example, increasing the buffer size from 1 to 2 has increased the accuracy by almost %24 (%35.69 – %11.35). Adding buffer size from 2 to 3 also improved the accuracy by around %6.

However, it is well known that the buffers have a limited contribution to the performance of the production system beyond a certain point. In the aforementioned case for example (case 8), after buffers size 54, almost no significant improvement is observed, and after buffer size 162 the results improvement almost stops. The minimum cycle time obtained by approximation for this case is almost 44838 seconds, corresponding to the buffer size of 166. The addition of further buffer space beyond this point has little impact on the reduction of cycle time.

4.4 Summary:

In this chapter a numerical study has been performed to validate the proposed approximation, examine the accuracy, and to study the effects of several different factors on the strength of the approximation.

At the beginning, several cases have been considered under different scenarios and the approximation procedure based on the models' parameters are performed. Also, for the purpose of comparison, the studied systems are modeled in simulation software (ProModel), simulated corresponding to each buffer size, and the comparison of results are presented.

Interpretations and analysis of the observation of numerical results are given in section 3. This analysis helps further in illustrating the strength and limitations of the proposed method based on variety of factors such as Mean Time to Failure, Mean Time to Repair, batch size, and buffer size.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

This thesis deals with the problem of performance evaluation in a multi-product production line with unreliable workstations. The few existing related solution approaches found in the literature mostly concentrate on either reliable workstations, or a single-product production line. To the best of author's knowledge, there is no analytical solution for problems that have similar characteristics to the one defined in this thesis. This thesis proposes an analytical approximation method to evaluate the cycle time as a performance measure in multi product production lines where workstations are unreliable.

Due to the complexity of the multi-product system in terms of defining a stable performance measure that does not change when a product type changes, and for the purpose of feasibility and practicality of the defined problem, we proposed generalization of production rate which is a key parameter to evaluate Total Cycle Time. The proposed general production rate would be independent of changes in product type. As a part of the generalization process, we modelled the variety of processing times in each work station by a generally distributed processing time. This will lead to obtain the desired general performance measures such as production rate and throughput time that enables us to calculate the Total Cycle time as the ultimate performance measure. GI/G/1/N queuing model is applied to a modified decomposition algorithm in order to obtain the general production rate as the basic measure of performance.

The strength and accuracy of the proposed approximation method is examined through a numerical study presented in chapter 4. Several cases and different scenarios are studied in order to explore areas of strength and limitations of the proposed method. The accuracy of the method is evaluated and an analysis of results is performed in order to give the best explanation of the behaviour of approximation under different scenarios.

The studied cases include cases with 5 and 10 workstations in series, 2 and 5 product types, and deterministic or generally distributed processing times. It has been observed that Mean Time to Failure and Repair (MTTF and MTTR) has a significant effect on the accuracy of the proposed approximation method. The reason for such an impact is the significance of these two parameters in the calculation of the coefficient of variation of the completion time.

Effects of buffer size were also studied for all cases. By performing the experiments for a considerably large variety of buffer sizes, the contribution of buffer in the performance of the systems and in accuracy of the proposed method were examined. By increasing buffer capacity the accuracy and performance of the line is improved up to a certain point, and thereafter, increasing buffer size will not make a significant improvement.

Another result obtained from the numerical study is increasing the strength of approximation when product type increases. Increasing product types will make the studied system closer to a system with generally distributed processing times, and since we are using the concept of general distribution for tackling the variety in product type, the observed improvement in the approximation is well explained. The effect of batch size on the power approximation has also been examined and a direct relationship between batch size and the method's accuracy were observed.

5.2 Future work recommendations

The proposed method and consequent numerical study in this thesis creates and establishes a new path for researchers in the area of performance analysis in manufacturing systems. It fills the gap between theory and practice when it comes to monitoring and evaluating the performance in smaller business where the complexity of the production system is considerably high. The idea of approximating product variety by general distribution of processing times that was implemented in this thesis is a new concept and has a large place for improvements and more elaborations.

Among possible enhancements, consideration of queuing models with group arrival and integrating them into the approximation model of this thesis can be investigated. As more elaboration, examining a wide range of statistical distributions of processing time (such as Normal, Erlang, or Exponential distributions) corresponding to each product type can be considered.

Moreover, the methods developed so far are designed for serial production lines. However, the use of series-parallel systems is also common in industries. A series-parallel production system is a production line where each work station consists of more than one machine in parallel. The reasons for adding machines in parallel are achieving higher production rate and greater reliability (Dallery and Gershwin, 1992). There is no such analytical method that could be directly used for the evaluation of a series-parallel production system that produces several product types. So, there is a need for extending the approximation method proposed in this thesis to series-parallel systems.

REFERENCES

1. Abdul-Kader, W., and Gharbi, A., 2002, Capacity estimation of a multi-product unreliable production line, *International Journal of Production Research,* 40 (18), 4815-4834.

2. Abdul-Kader, W., 2006, Capacity improvement of an unreliable production line - an analytical approach, *Computers and Operations Research,* 33 (6), 1695-1712.

3. Aized, T., Takahashi, K., 2007, Modeling and performance optimization of multiple product FMS using colored Petri net and response surface methods, International Symposium on Assembly and Manufacturing, Ann Arbor, Michigan, USA, July 22-25.

4. Albores_Velasco, F.J., and Tajonar_Sanabria, F.S., 2004, Analysis of the GI/MSP/c/r queuing System, *Non-English name*, 4 (1), 46-57.

5. Alexandros, D.C., and Chrissoleon, T.P., 2008, Exact analysis of a two-workstation one-buffer flow line with parallel unreliable machines, *European Journal of Operational Research*, 197 (2), 572-580.

6. Askin, R. G., and Standridge, C. R., 1993, *Modeling and Analysis of Manufacturing Systems*. John Wiley & Sons, Inc., NY.

7. Asmussen, S., and Moller, J.R., 2001, Calculation of the steady state waiting time distribution in GI/PH/c and MAP/PH/c queues, *Queuing Systems*, 37 (1-3), 9-29.

8. Aziz, A., Jarrahi, F., Abdul-Kader, W., 2010, Modeling and Performance Evaluation of a Series-Parallel Flow Line System with Finite Buffers, *INFORE*, 48 (2), 103-120.

9. Baker, K.R., Powell, S.G., and Pyke, D.F., 1994, A predictive model for the throughput of unbalanced, unbuffered three-station serial lines, *IIE Transactions*, 26 (4), 62-71.

10. Banks, J., Carson, S.J., Nelson, L.B., and Nicol, D.M., 2005, Discrete-Event System Simulation, Pearson Prentice Hall, Inc., NJ.

11. Bard, Y., 1974, Non-linear parameter estimation, Academic Press, NY.

12. Blumenfeld, D.E., 1990, A simple formula for estimating throughput of serial production lines with variable processing times and limited buffer capacity, *International Journal of Production Research*, 28 (6), 1163-1182.

13. Borelli, F., Falcone, P., and Vecchio, D.C., 2007, Event-based receding horizon control for two-stage multi-product, *Control Engineering Practice,* 15 (12), 1556–1568.

14. Buzacott, J.A., 1971, The role of inventory banks in flow-line production systems, *International Journal of Production Research,* 9 (4), 425-436.

15. Buzacott, John A., and Shanthikumar, G., 1993, Stochastic Modeling of Manufacturing Systems, Prentic-Hall, Inc., Cliffs, NJ.

16. Choudhury, G., 1998, On a batch arrival Poisson queue with a random setup time and vacation period, *Computers & Operations Research*, 25 (12), 1013-1026.

17. Conway, R., Maxwell, W., McClain, J.O., and Thomas, L. J., 1988, The role of work-in-process inventory in serial production lines, *Operations Research,* 36 (2), 229-241.

18. Dallery, Y., David, R., and Xie, X. L., 1988, An efficient algorithm for analysis of transfer lines with unreliable machines and finite buffers, *IIE Transactions,* 20 (3), 280-283.

19. Dallery, Y., and Gershwin, S.B., 1992, Manufacturing flow line systems: a review of models and analytical results, *Queueing Systems Theory and Applications*, 12 (1-2), 3-94.

20. Gaver, D.P., 1962, A waiting line with interrupted service including priorities, *Journal of the Royal Statistical Society,* 24 (1), 73-90.

21. Gershwin, S.B., 1987a, An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking, *Operations Research*, 35 (2), 291-305.

22. Gershwin, S.B., 1987b, Representation analysis of transfer lines with machines that have different processing rates, *Annals of Operations Research*, 9 (1), 511-530.

23. Gupta, U.C., and Banik, A.D., 2006, Complete analysis of finite and infinite buffer GI/MSP/1 queue—A computational approach, *Operations Research Letters*, 35 (2), 273-280.

24. Haskose, A., Kingsman, B.G., and Worthington, D., 2004, Performance analysis of make-to-order manufacturing systems under different workload control regimes, *International Journal of Production Economics*, 90 (2), 169-186.

25. Hillier, F.S., and So, K.C., 1991, The effect of machine breakdowns and interstage storage on the performance of production line systems, *International Journal of Production Research,* 29 (10), 2043-2055.

26. Johri, P.K., 1987, Linear programming approach to capacity estimation of automated production lines with finite buffers, *International journal of production research*, 25 (6), 851-867.

27. Ke, Jau-Chuan., 2007, Batch arrival queues under vacation policies with server breakdowns and start-up /closedown times, *Applied Mathematical Modelling*, 31 (7), 1282-1292.

28. Kendall, D.G., 1953, Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of Imbedded Markov Chains, *Annuals of Mathematical Statistics*, 24, 338-354.

29. Kenne, J.P., and Gharbi, A., 2004, A simulation optimization based control policy for failure prone one-machine, two-product manufacturing systems, *Computers and Industrial Engineering,* 46 (2), 285–292.

30. Omar, M.K., and Teo, S.C., 2007, Hierarchical production planning and scheduling in a multi-product, batch process environment, *International Journal of Production Research*, 45 (5), 1029-1047.

31. Powell, S.G., and Pyke, D.F., 1996, Allocation of buffers to serial production lines with bottlenecks, *IIE Transactions,* 28 (1), 18-29.

32. Tempelmeier, H., and Burger, M., 2001, Performance evaluation of unbalanced flow lines with general distributed processing times, failures and imperfect production, *IIE Transactions*, 33 (4), 293-302.

33. Xu, X., and Zhang Z.G., 2006, Analysis of multi-server queue with a single vacation (e, d)-policy, *Performance Evaluation*, 63 (8), 825-838.

34. Zipkin, P.H., 1995, Performance analysis of a multi-item production-inventory system under alternative policies, *Management Science*, 41 (4), 690-703.

VITA AUCTORIS

NAME:	Farshad Jarrahi
PLACE OF BIRTH:	Tehran, Iran
YEAR OF BIRTH:	1975
EDUCATION:	Azad University, South Tehran Branch
	Tehran, Iran
	B.Sc. in Industrial Engineering