Politecnico di Milano
Dipartimento di Meccanica
Dottorato di Ricerca in Ingegneria Meccanica

Analysis of the Output Variability in Multi-Stage Manufacturing Systems

Tesi di dottorato di:
Ramiz Assaf

Relatore:
Dr. Marcello Colledani
Tutore:
Prof. Tullio Tolio
Coordinatore del programma di dottorato:
Prof. Gianpiero Mastinu

2011 — XXIV ciclo

Politecnico di Milano
Dipartimento di Meccanica
Via La Masa 1, 20156 — Milano

Politecnico di Milano
Dipartimento di Meccanica
Dottorato di Ricerca in Ingegneria Meccanica

Analysis of the Output Variability in Multi-Stage Manufacturing Systems

Doctoral Dissertation of:
Ramiz Assaf

Advisor:
Dr. Marcello Colledani
Tutor:
Prof. Tullio Tolio
Supervisor of the Doctoral Program:
Prof. Gianpiero Mastinu

2011 — XXIV cycle

To my beloved Family for their endless support.
To my fiancée for her patience and love.
To my friends for being there for me.

Acknowledgments

I would like to thank my advisor Dr. Marcello Colledani for teaching and guiding me during my study period. I would also like to thank Dr. Andrea Matta who was a second supervisor for me. Furthermore, I would like to thank all my colleagues at Politecnico di Milano who created an amazing atmosphere during my study period.

Abstract

Manufacturing System Engineering methods have been developed in the last decades for investigating the dynamic behavior of manufacturing systems, for estimating their performance and for supporting their efficient design, improvement and reconfiguration. Typically, these approaches are focused on the first order performance measures of manufacturing systems, such as the average throughput, the average Work in Progress (WIP) and the average lead time. During the system configuration / reconfiguration phase, these tools are used to select system solutions that profitably exploit the trade-offs between these first order performance measures. Higher order performance measures are generally difficult to analyze and are rarely considered in these analyses.

However, in the presence of random events and disturbances in the production, higher order performance measures are relevant to correctly predict the system output. Indeed, due to the production variability, the observed performance can be highly different from the average performance. This problem may directly corrupt the profitability of those systems designed only by considering the mean performance of the system that are not robust to disturbances. Typical sources of variability in the production system behavior are random failures occurrences and durations. A real case in the automotive sector reports that the daily output of the production system composed of 22 machines affected by the occurrences of 144 different failure modes, has a coefficient of variation, estimated from the available field data of four months, equal to 0.263. Thus, it is highly probable that the weekly demand will not be met if the system is designed only considering its average performance.

Pioneer works in the field on output variability evaluation, proposed the output asymptotic variance rate, as a relevant performance measure. The output asymptotic variance rate is defined as the limit of the variance of the output process per unit time, as time approaches infinity. This performance measure can be used as a machine parameter to compare different machines during the system design phase. Moreover, it can be also used in combination with the average production rate to calculate the service level of the manufacturing system.

In spite of the industrial relevance of this problem, the number of papers discussing the output variability in production lines is limited. Moreover, general and efficient methods to estimate the variance of long multi-stage systems with finite buffer capacities and unreliable machines do not exist. Thus, very little is known on how to manage production systems to reduce the variability of their output. The proposed research aims at developing new methods to estimate the output variability in multi-stage manufacturing systems with finite buffers, by using approximate analytical methods.

The first part of the work is focused on the exact analytical evaluation of the output variability of small production systems modeled as discrete time – discrete state Markov chains with binary reward. On the one hand, the exact analytical evaluation allows understanding the behavior of the output variability as a function of the system parameters. On the other hand, by enabling to study simple two-machine one-buffer systems, it serves as a building block to the aggregation/decomposition methods developed in this thesis for the analysis of long manufacturing lines.

The reminder of the work is focused on the evaluation of the output variability for long multi-stage manufacturing lines, by using approximate analytical methods based on decomposition and aggregation techniques. Three approaches are developed and extensively compared.

The first approach uses the traditional decomposition method for throughput estimation in multi-stage systems. The manufacturing system is decomposed into two machine one buffer subsystems. The idea of this method is to create building blocks that conserve the average material flow throughout the system. Upon convergence, the asymptotic variance rate is calculated on the last building block in the system, using the exact analytical method developed in this thesis.

Although this approach shows good accuracy towards simulation in the estimation of the mean throughput, it performs poorly while calculating the output asymptotic variance rate of the system (average error is about 34%). A modification of this approach is then introduced to consider a more complex three machines - two buffer subsystem as the last building block. This results in lower errors in calculating the output asymptotic variance rate (average error is about 5.25%). However this modified approach is highly computational intensive in general and only applicable when the last two buffers have small sizes.

The second approach developed for the calculation of the output variability for multi-stage manufacturing lines is based on the aggregation technique. The method considers a first sub-system formed by the first machine/buffer/machine dipole in the system, and it calculates the average throughput and the asymptotic variance rate for this small system. Then it approximates the behavior of this sub-system to the behavior of a geometric pseudo-machine with one failure mode having the same asymptotic throughput and variance rate. Afterwards, a new sub-system is evaluated formed by this pseudo-machine and the following buffer/machine of the original line. This process is repeated until the last machine of the manufacturing system is included. This method reduces the error in calculating output asymptotic variance rate to around 8%. However it suffers of low accuracy in the calculation of the average throughput, as it does not consider the effect of blocking propagation properly.

The third approach proposed in this thesis is a decomposition method that conserves the asymptotic variance rate throughout the system's stages. New decomposition equations are developed that propagate both the average throughput and the variance rate throughout the system. This approach shows very small errors for both the average throughput and the asymptotic variance rate. Also, this method shows to be fast, enabling to evaluate a ten machine line with relatively large buffer sizes within a minute.

Based on the developed methods, a new approach for the optimization of manufacturing systems has been proposed. This new approach aims at providing the optimal configuration of the buffers that minimize the output variability using the non linear conjugate gradient optimization. The proposed approach is used for the analysis of an industrial case featuring a buffered multi-stage manufacturing system. The results provided by the solution of this optimization problem shows that optimal solutions that only consider the average throughput of the system are sub-performing if the variance rate of the system is also considered, thus paving the way to the development of new approaches for designing the system to meet target service levels. Finally, this research provides guidelines for production managers to handle and reduce the output variability in their manufacturing systems.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The Evolution of Manufacturing Systems

Manufacturing has evolved over the course of history of man-kind from skilled single workers to complex, fully automated and computerized systems. Most of the development in manufacturing systems, as well as related fields, occurred in the 20th century.

Advances in manufacturing systems have happened on different levels. There has been a continuous development of conventional and unconventional processing technologies paralleled with developments in system configurations, automation and control. However, the most important developments in the realm of manufacturing were at the philosophical level.

Pioneers such as Taylor, Ford and Deming came out with new manufacturing system philosophies. Their ideas caused a paradigm shift in the industry. For example, prior to Deming, mass production was the goal of all factory owners, practitioners, engineers and even researchers. Deming's ideas and philosophies shifted the eyes of the industry towards quality.

These new philosophies and paradigms have caused change in the strategies, goals, methodologies, approaches, and even support tools of manufacturing systems [26]. Furthermore, manufacturing systems develop, interact and overlap with other domains or fields. For example, information technology (IT) made an enormous impact on the way manufacturing systems' data is collected, summarized and analyzed. IT also created connections between the different parts of an organization, enabled planning at an organizational level, and created holistic planning approaches such as the Enterprise Resource Planning (ERP) [68].

In current times, it is easy to list a wide range of manufacturing philosophies, such as "agile and flexible manufacturing systems", "lean production and manufacturing", "customer driven intelligent manufacturing systems" [70] and "the integration of products, processes and production systems" [86]. Most of the new paradigms are driven by either a service and value mission, global integration / competition and/or information sharing.

Due to the effect of open markets and globalization, a huge emphasis is being given to customer oriented processes within an organization. This process starts with marketing, followed by design and manufacturing, and ending up with logistics. [46, 27]. Also, organizations need to know the customers' expectations, and how to adapt to meet those expectations.

1.2 Performance Measures During the Evolution of Production Systems

As manufacturing systems evolve, their performance measure evolves as well. During Taylor's time, a huge driving force for manufacturing systems was the economies of scale, as explained previously. The ideas Taylor came up with match the goal of producing as much as possible with a minimum work force. Methods like time studies were used extensively. The performances measured during that era represented "how much time can I save?" or "how much improvement I can get by changing the way things are done?". In the quality era (if one can call it so) new questions came up, such as "How many defect-free products can our system produce?" [see [67] and [73]].

Nowadays, arising questions like "How and when does my customer want the product?" or from a design point of view "Does my product satisfy my customers' needs?" and "How fast can I send the products to my customer?". The examples are infinite, and it should be realized that new philosophies build up on previous ones, only changing the main goal needs .

It should be clear at this point that if the customers are the center of focus, producers need to build strong and reliable business connections with them based on pure trust. Products sent to them should be of a known quality, be reliable, and should be delivered *consistently*, with no delays. Relevant performance measures for customer oriented manufacturing systems and organizations, for example, is the Service Level (SL) which can be defined as the probability to meet a customer order of size x within a certain time window t [79]. The evaluation of the SL for a manufacturing system requires the calculation of the total production's first two moments. Since the first moments have been studied a lot, the missing piece is the second moment, i.e the output variability which is the focus of this thesis.

1.3 The Design and the Evaluation of Manufacturing Systems

The design of manufacturing systems is a complex task that needs considerable efforts and a huge amount of resources. The output of such an endeavour is a manufacturing system that meets certain characteristics, within time and budget constraints see [58]. On the one hand, a manufacturing system should be capable of realizing products needed with certain quality, cost and complexity characteristics. On the other hand, the production system used should satisfy a certain criteria, such as setup times, percentage up time and productivity. Nowadays, even more dimensions affect the design process, e.g. energy consumption, machine/product life cycle, sustainability issues and finally market dynamics [86].

Market dynamics is a very important factor that can change the typology of a designed manufacturing system. For example, a more customized market could need the use of cell manufacturing systems that can adapt to high customization. Another example could be a highly competitive market with strict timing requirements. Any loss of sale could mean the loss of a customer. Thus, reliability should be the main factor in the system's design.

Manufacturing systems output variability can affect greatly companies' performance. The higher the output variability the lower the chances are to satisfy certain scheduled order, and this could incur them extra cost. The form of these cost lies in lost orders, penalties, reputation and/or customer compensations from one hand, or using high safety stocks from another hand. These extra costs are not needed if the manufacturing system is more stable and producing less output variability.

Manufacturing systems designers, plant, production and logistics managers who consider the average production rate as their main performance measure to optimize, are neglecting a great deal of the problem. For example, in practice contracts are done considering a certain service level, which in fact considers the output variability, indirectly. The problem arises when planning to measure something you have no control over.

1.4 How Is Output Variability Generated?

Output variability is generated due to randomness of events and disturbances affecting the production system. Examples include:

- Process variability: Manufacturing contexts where manufacturing stages are not deterministic but random variables. Manual assembly operations are considered to have a high process variability. In assembly systems, operators' time to finish a task depends on a lot of factors like motivation, concentration, fatigue...[56, 74].
- Failures in machines: In a real manufacturing system, machines are not reliable; different failure modes can happen, and they happen randomly. Since machines are composed of many complex components, machines can fail in multiple failure modes and each mode with a different severity and frequency.
- Repair times: When a machine fails the time needed to repair is not deterministic due to many reasons. First of all, there is an identification time i.e. the time needed to know that a machine has failed and to identify what is the nature of the failure. Then,

there is the actual repairing process which takes most of the time, and finally the calibration and setup times.

- Suppliers' Delivery Variability: In principle, if suppliers have very high delivery variability, they will impact the output variability of the production systems as this increases the variability input to the system. The same is applied to repair or exchange parts.

- Quality perspective: If a machine produces defective products, the production engineers will need to stop the production and fix the machine, recalibrate, and setup before continuing production. This process could be identified as a quality failure of the machine. Furthermore, having buffers in the production line could cause a delay in detecting quality problems causing more products being discarded [21].

The scope of this work is to consider the output variability generated from the failure/repair mechanism happening to the machines of the manufacturing system. Hence, we assume there is no delay in raw material nor in repair parts, and machines produce no defects.

1.5 Output Variability Performance Measures

The interest of this work is to give a general method that calculates different output variability measures for production systems. The main variable of interest for this thesis is the total production output during the time period $[1, t]$. This variable, denoted with Z_t, is random and its first two moments are the expected value $\mathbb{E}[Z_t]$ and the variance $var[Z_t]$, which both depend on the time period of evaluation as they are increasing functions in t. The coefficient of variation of Z_t is given by:

$$cv[Z_t] = \frac{\sqrt{var[Z_t]}}{\mathbb{E}[Z_t]} \tag{1.1}$$

and it is a decreasing function in t.

Miltenburg [61] and Tan [79] have shown that for large t (that approaches *infinity*) Z_t is approximately normally distributed with a mean $e{\cdot}t$ and a variance $v{\cdot}t$, where e is the mean production rate or the mean throughput of the system, and v is the asymptotic variance rate. Such asymptotic measures can be calculated as:

$$e = \lim_{t\to\infty} \frac{\mathbb{E}[Z_t]}{t} \tag{1.2}$$

$$v = \lim_{t \to \infty} \frac{var[Z_t]}{t} \tag{1.3}$$

e and v do not depend on the time, so they can be used as output characteristic parameters during the system design phase. The coefficient of variation can be written in terms of e and v as:

$$cv[Z_t] \approx \frac{\sqrt{v}}{e\sqrt{t}} \tag{1.4}$$

which approaches *zero* as t approaches infinity, i.e., as the time increases the dispersion of Z_t decreases.

A long term measure of the relative output variability is the index of dispersion [4]:

$$d = \frac{v}{e} \tag{1.5}$$

Finally, approximating Z_t with a normal distribution allows the calculation of the system service level (SL), defined as the probability to meet a certain customer order (composed of x products) within a certain time t [79]:

$$SL(x,t) \approx P[Z_t \geq x] = 1 - \Phi\left(\frac{x - e \cdot t}{\sqrt{v \cdot t}}\right) \tag{1.6}$$

where $\Phi(\cdot)$ is the cumulative normal distribution function. The SL can also be written in terms of the error function as:

$$SL(x,t) = P[Z_t \geq x] = \frac{1}{2}\left(1 + erf\left(\frac{e \cdot t - x}{\sqrt{2v \cdot t}}\right)\right) \tag{1.7}$$

where $erf(\cdot)$ is the error function.

1.6 The Importance of the Output Variability - Real Industrial Cases

The significance of the output variability was demonstrated by Gershwin [29] using simulation of a representative line. His experiments showed that the variability of the output is very high and the standard deviation can be around 10%. Tan [81] collected data from a production line during a period of three months and arrived to same result of Gershwin, concluding that considering the steady-state behavior of a production line is not adequate to design and control production lines in rapidly changing, dynamic environment.

In this section, we analyze date collected from two real manufacturing systems, showing the importance of considering the output variability. Afterwards, we show how much impact reducing the output variability has on the service level of the manufacturing system. The first case is Scania's Engine-Block production line, studied in depth in [18] and [20]. While the second case will be named "Case 2" due to confidentiality agreement with the company.

1.6.1 Case 1 - Scania

Scania has a production line consisting of 22 working stations with a total of 144 failure modes. The layout of the system can be seen in Figure 3.3. Figure 1.1 depicts the daily production data collected during the study period of 102 days. The observed daily production is quite variable. Furthermore, Figure 1.2 shows the value of $cv[Z_t]$ calculated over a time period (t). The value of $cv[Z_t]$ can be calculated from equation (1.1), for example, for a certain time window (t) e.g. (t) equals 5 days, we transform the daily production series into a cumulated 5-day production series, then cv calculated from this new series is $cv[Z_5]$.

Figure 1.2 depicts the plot of $cv[Z_t]$ for t ranging from 1 to 30 days. It should be mentioned that as the original data consists of 102 data points, a problem arises when reaching high t values as the plotted cv will correspond to only few data points, which decreases the accuracy of estimating cv_t. Nevertheless, the results show that although cv decreases rapidly with increasing t, at a time window $t = 30$ days cv is not equal to zero.

The service level of the production line of Scania can be calculated for a range of demand quantities x at a certain time t. For example, from Figure 1.2 $cv[Z_{t=10}]$ is equal to 0.115, $\mathbb{E}[Z_{t=10}] = 1329$ products, and the asymptotic variance rate estimated using equation (1.3), is 12.1 $^{\text{parts}^2}/_{\text{10days}}$. The service level of this example is plotted in Figure 1.3.

Now, what happens to the SL, if we were able to reduce the asymptotic variance rate V of the output manufacturing system by 35%? If we succeeded, the new asymptotic variance rate will be equal to 7 $^{\text{parts}^2}/_{\text{10days}}$. Now for each $\mathbb{E}[Z_t]$ we obtain an improvement in the SL, as shown in Figure 1.4. For example, assuming the shipment of 1175 products has a due date after 10 days, the old SL corresponds

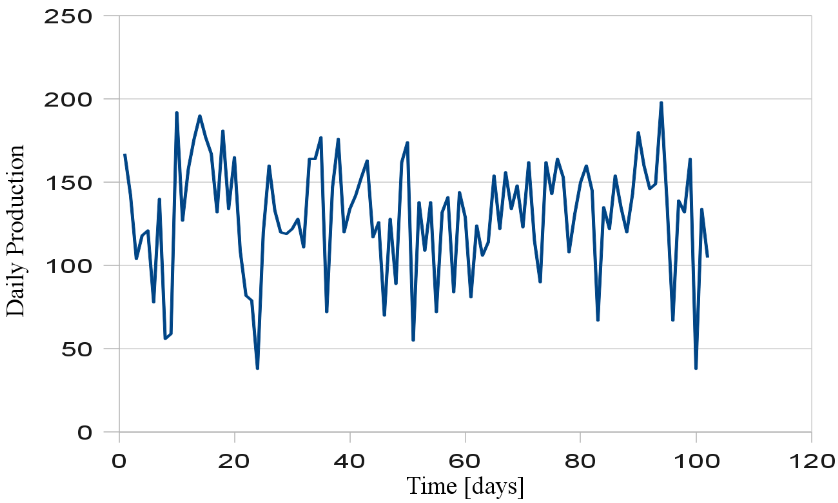


Figure 1.1: The daily production during the studied period

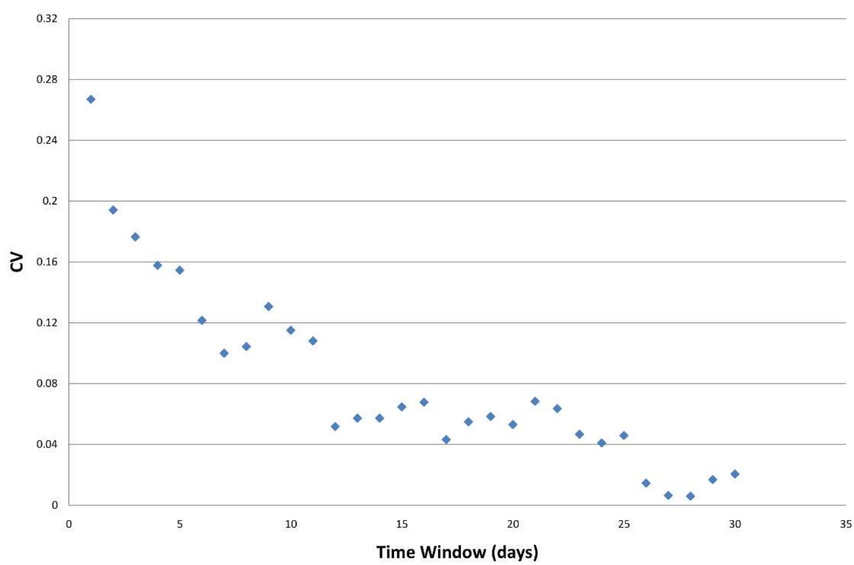


Figure 1.2: Scania Case - The estimated coefficient of variation at time t from the production during the observed period

to 85% while the new SL corresponds to around 90%. Thus the SL is highly affected by the asymptotic variance rate.

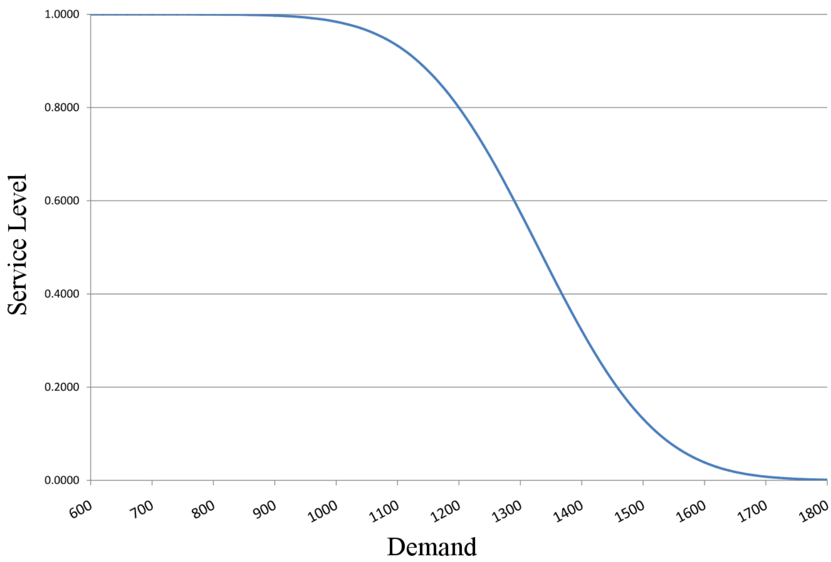


Figure 1.3: Scania Case - The service level of the system as a function of the demand at t = 10 days

1.6.2 Case 2- Car Parts Producer

This company has a multi-station line with 28 machines, with a set of failure modes. Production data were collected during 244 days. Figure 1.5 shows the daily production for this case, while Figure 1.6 shows the values of $cv[Z_t]$. It is shown from 1.6 that $cv[Z_t]$ starts from the value of 0.17 when t is equal to 1 day, and decreases slowly to finish around 0.072 when t =30. So the $cv[Z_t]$ in this case did not reach zero. For $t \rightarrow$30 days surprisingly, it seems that it stabilizes around 0.07.

Output variability impact on the service level- Case 2

The SL of the system can be calculated, as in the Scania case, over a range of demand quantities x arriving after a time t=14 days. From Figure 1.6 the cv for t = 14 days is equal to 0.09, $\mathbb{E}[Z_t]$ = 16732

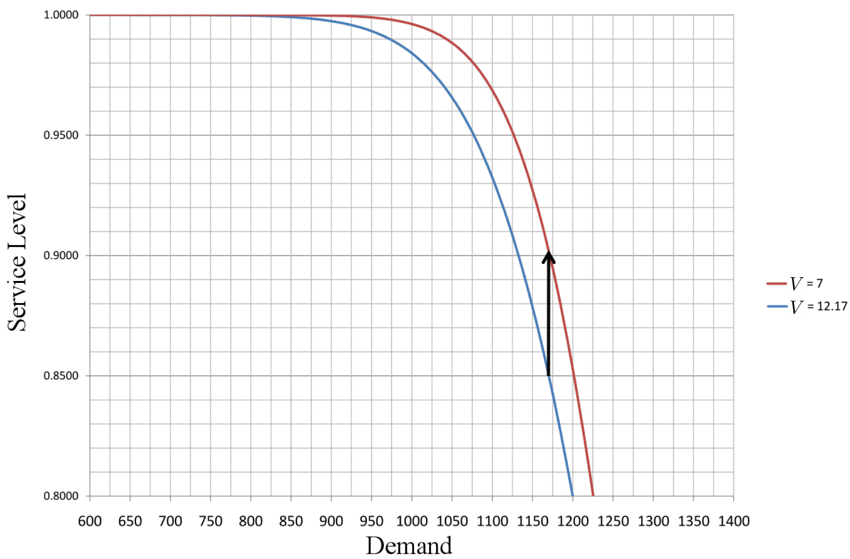


Figure 1.4: Scania Case - The service level of the system at t=10 days, for two different values of v 12.1 and 7 parts$^2/_{10days}$

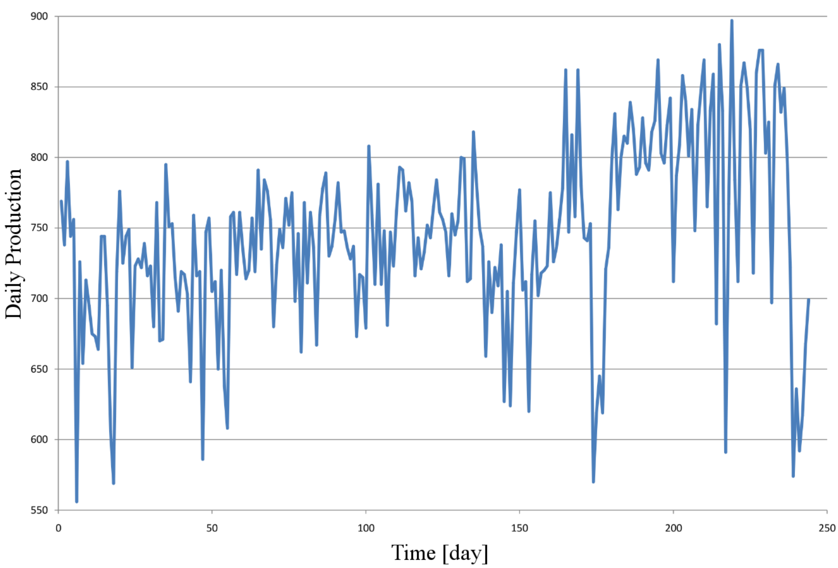


Figure 1.5: Case2 - The daily production during the studied period

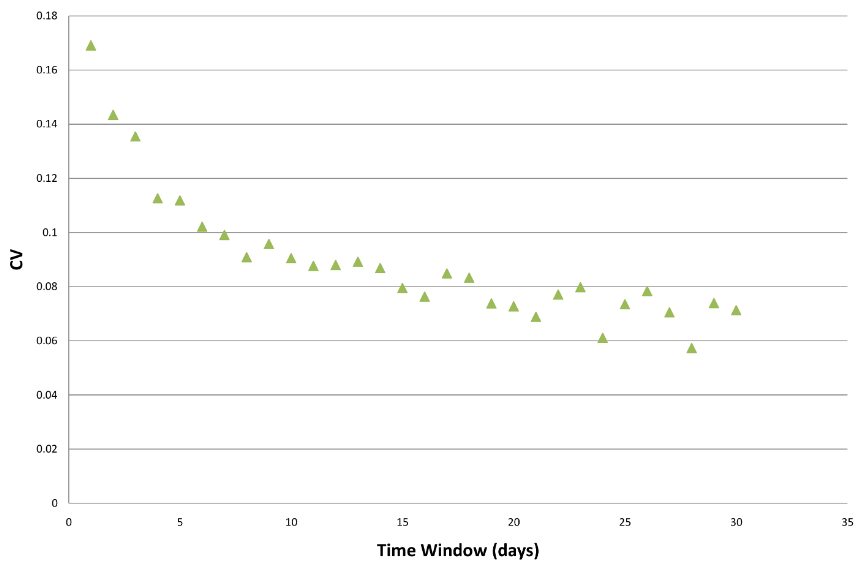


Figure 1.6: Case 2 - The estimated coefficient of variation at time t from the production during the observed period

products and the asymptotic variance rate, calculated using equation (1.3), is 144 $^{\text{products}^2}/_{\text{14days}}$. The SL of such example is plotted in Figure 1.7.

Figure 1.8 depicts the effect of decreasing V on the service level following the same procedure used in the Scania case. For example, a demand x that is equal to 15000 products and to be delivered in two weeks ($t = 14$ days) has a SL that is equal to 84.5%. If the production system was improved to produce a lower asymptotic variance rate of 120 or of 95 $^{\text{products}^2}/_{t}$, the SL would increase to 87% and 89.5%, respectively.

1.7 Research Objectives

The first objective of this work is to develop an analytical method for the analysis of output variability in manufacturing systems. The developed method should be more general than previously proposed methods and approaches in terms of the complexity of the machine structure considered.

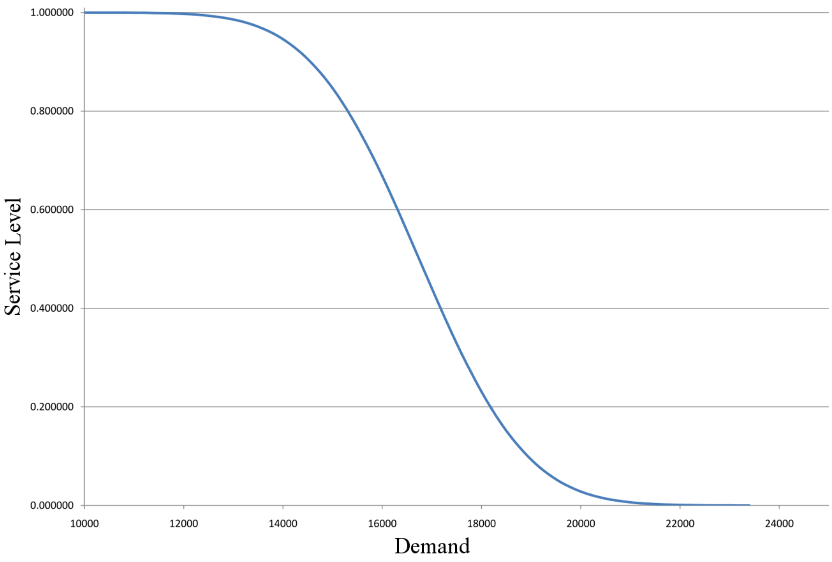


Figure 1.7: Case 2 - The service level of the system as a function of the demand at t = 14 days

The second objective is to be able to give an explanation to the behavior of output variability, and its dependency on the manufacturing systems' parameters. The output variability's main performance measure (i.e. the asymptotic variance rate) was previously shown to have a non-monotonic behavior as a function of buffer capacities (for systems with buffers) [[11] and [82]]. So far, no complete understanding exists for such behavior.

Finally the last objective of the work is to provide tools, guidelines, and rules that would help practitioners to reduce the output variability in multi-stage manufacturing systems.

1.8 Outline of the Thesis

The remainder of the thesis is organized as follows:

Chapter 2 - Literature Review: the first part presents a comparison between simulation and analytical models for the evaluation of

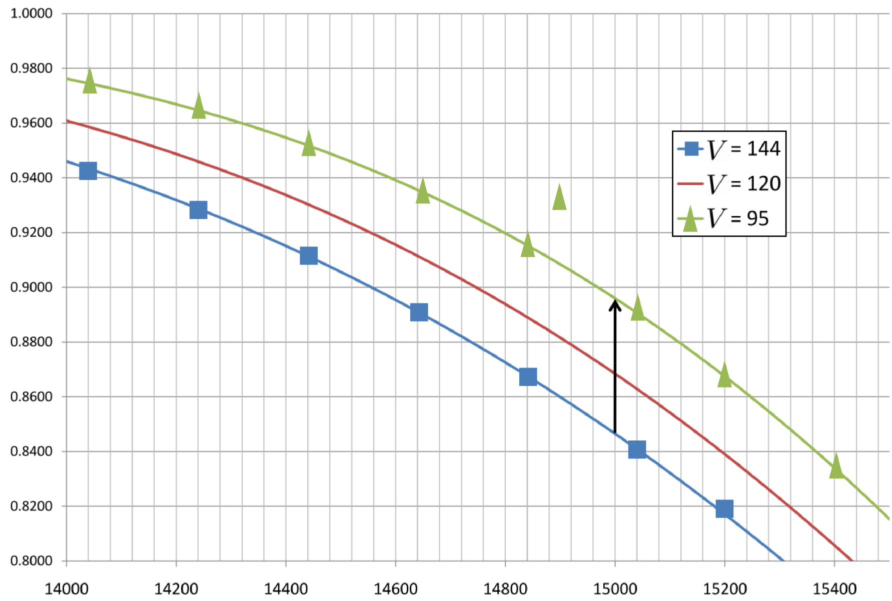


Figure 1.8: Case 2 - The service level of the system at t=14 days, for different values of asymptotic variance rate

manufacturing systems. Then, main works for exact and approximate analytical evaluation of average performance measures are presented. The second part presents a deep analysis of the works presented concerning the evaluation of the second moments of the output i.e. The output variance, the asymptotic variance rate, transient behavior of the system and the inter-departure distribution of the output.

Chapter 3 - Production System Model: the taxonomy used to characterize the behavior of the production system and to represent it in a formal way is proposed. Moreover, the production model considered in this thesis is defined and the main modeling assumptions are introduced.

Chapter 4 - The Analysis of The Output Variability in Small Production Systems: methods for the exact evaluation of the output variability of General Markovian isolated machines, two machine lines and small production systems are presented. Furthermore, the impact of changing machine parameters and the buffer capacity on the output variability is investigated.

Chapter 5 - The Analysis of Output Variability in Long Multi-Stage Production Systems: a set of different approximate analytical methods (CMT, CMT3, AC, $\hat{A}$C and AGG) that can be used for the evaluation of long multi stage production lines that were developed in this work are presented.

Chapter 6 - System Behavior: The comparison among the different proposed approaches for the evaluation of long production lines in terms of accuracy and speed is carried out, giving guidelines on when to apply each method. Furthermore, an analysis of the factorial design using Design of Experiments (DOE) approach using ANOVA to identify factors that affect the accuracy of the proposed methods. This chapter also presents a study of the buffers capacities effect on the output asymptotic variance rate.

Chapter 7 - The Optimization of Production Systems Considering the Output Variability: the optimization of average measures of the output performances always ignores the impact of output variability, in this chapter, a new optimization approach that uses the Conjugate Gradient method that considers the output variability and service level was developed, formalized and tested, giving examples on different cases.

Chapter 8 - The Analysis of Levissima Water Bottling Production System: This chapter presents a case study, in which the developed methods proposed in chapters 5 were used for the analysis phase, and the optimization procedure developed in chapter 7 has been applied to minimize the output variability in the production line. It also gives guidelines on how to evaluate and minimize the output variability of an existing production system.

Chapter 9 - Conclusion and Recommendations: In this chapter managerial rules for reducing the output variability will be given. Then a summary of the main conclusions of the research, extensions and future work topics will be presented.

Chapter 2

Literature Review

In this chapter a review on methods for the evaluation of production systems is provided. The main analytical methods for the performance evaluation of serial production lines in terms of output average and variability performances is also reported. It is also necessary to introduce the contributions done as constitute the an important basis for the approaches proposed in this thesis.

2.1 Performance Evaluation of Production Lines: Simulation vs. Analytical Methods

Research in the area of the evaluation of production systems has been divided into two distinct areas: simulation method and analytical evaluation. Simulation method involves the representation of the real manufacturing system in a computer-based model via the use of an appropriate simulation package. Simulation models are capable of handling complex model structures, however they need validation to be representative of the true manufacturing system. On the another hand analytical methods involve formal mathematical solution to the problems. Due to the complexity involved two approaches are used namely, exact and approximate methods. Exact analytical solutions are feasible for simplified models, and usually small scale problems. Approximate methods derive approximate solutions often by means of appropriate and efficient algorithms to actual mathematical problem

Papadopoulos et al. [70].

2.2 Analytical Method for Average Performances Evaluation

2.2.1 Exact Analytical Methods

The use of exact analytical techniques for average evaluation of manufacturing system is mainly based on queuing theory. Reviews on application of the queuing theory in manufacturing can be found in [38, 71, 10] for general queuing models. The problem with these approaches is that when the system increases in size (by adding buffers of more machine states), the time and memory needed for evaluations grows exponentially. This fact limits the use of such methods for production systems, and lines evaluated are usually small [23, 29, 70].

2.2.2 Approximate Analytical Methods

The aim of approximate analytical methods is to relax some of the most limiting assumptions of the queuing theory to be able to model and evaluate a large set of real systems. Since Chapter 5 presents two approximate analytical methods based on the decomposition technique, the historical background of this approach is summarized.

A review of the most important works done in this area can be found in [23], and references found therein. The evaluation of the average throughput of the system was based on the work of [9] and obtained by averaging the production rate of each machine. Some works had the goal of demonstrating the properties of a production serial line by using the approximate analytical methods.

Gershwin and Schick [30] demonstrated the property of conservation of the average throughput in a production line. Muth [63], demonstrated the property of reversibility of a production line, i.e. inverting the order of the machines in the line, the average production rate remains constant. [35] proposed the first effective exact solution for a two-machine line, in which the Markov chain describing the behavior of the system is solved independently on the capacity of the buffer, following a product form solution.

Other types of problems have been analyzed next. Jafari and Shanthikumar [47] analyzed flow lines with imperfect and scrapped parts. Moreover, they extended the analysis to case of two machine lines with general uptime and downtime distributions. Yeralan and Muth [91] proposed a method in which repairing personnel were shared by different stations and the repair time depended on the availability of the operator. With the work of Gershwin [34] starts the series of works that more strongly influenced the next generation of researchers. He proposed a new decomposition approach based on the exact solution of the two-machine system already analyzed in Gershwin and Berman [35]. The model deals with discrete time assumptions, geometrically distributed failure and repair times, unique failure mode and finite buffer capacity. The proposed decomposition algorithm was later improved by the work of Dallery et al. [22]. New improvement of the approach followed. First of all, the consideration of multiple failure modes for each machine, i.e. the possibility that one machine can go down for different reasons and with different frequencies of failure and repair, was analyzed by Tolio et al. [88]. Moreover, in Le Bihan and Dallery [49] and Tan and Yeralan [83] new decomposition approaches were proposed.

Afterwards, the research has been directed toward an improved applicability of the proposed approach to real systems by dealing with the analysis of larger set of system architectures, by reducing the approximation error and by generalizing the methodologies assumptions. Firstly, assembly/disassembly systems have been considered [36, 13]. Later, systems characterized by non-linear flow of material were analyzed by Helber [41] and Diamantidis et al. [24]. Moreover system presenting closed loop architectures were studied by Gershwin and Werner [33], later research included systems considering multiple closed loops R. [72], Li et al. [53]. Multi-product systems have been recently considered both with queuing networks Baynat et al. [5] and Bitran and Tirupati [6] and with approximate analytical techniques by Altiok and Stidham [3] and Colledani et al. [16]. Moreover in Miltenburg [62] works analyzing the behavior of U-shaped lines are summarized. Recently new techniques for the evaluation of generally complex system layouts have been developed in Li [50], where the overlapping decomposition method is proposed. Colledani and Tolio [15] proposed the Two-level Decomposition which evaluates the performance of complex systems in which quality control is applied.

Production control policies for regulating the throughput rate of the system have been also studied with decomposition techniques. Relevant works in this area are the work of Gershwin [31] and Véricourt and Gershwin [89] for modeling and evaluating the performance of systems controlled by the Control Point Policy and the work of Bonvik et al. [7] which review and compare the performance of systems controlled under different policies. Finally, Matta et al. [59] analyzed the performance of assembly systems controlled with kanbans with the use of queuing networks. The variations of the kanban systems can be found in Junior and Filho [48].

2.3 Output Variability Evaluation

In spite of the relevance of this issue in industry, the number of papers discussing the variability of the output in production lines is fairly limited if compared to the papers on the prediction of first order performance measures of manufacturing systems. Moreover, the underlying assumptions of the available methods are over-simplistic, thus preventing their wide application in industry. Research contributions on production variability deal with both the cumulated production of a production line and the interdeparture times of the output process in a time interval.

The output variability of production lines was first studied by Miltenburg [61]. He proposed an exact numerical method to calculate the first two moments of the asymptotic measures of the output, i.e. the throughput and the asymptotic variance rate, that is the limiting variance of the output per time unit. The method considered small buffered production lines featuring unreliable machines with geometric/exponential failure and repair times. In addition he demonstrated that the method can model machines with different processing times. His approach is based on the state-space representation of the system and the use of the inverse of the fundamental matrix. Since the computational complexity of this method depends on the number of states modeling the system, only simple systems with small number of machines and buffer capacities can be analyzed with success.

Hendricks [42] presented an approach, based on the structural properties of Markov chains, to estimate the asymptotic variance rate of interdeparture times in production lines with exponential processing times, perfectly reliable machines and finite buffer capacities. This work was later extended by [43] to model machines with general processing times. He provided interesting insights on the role of the output autocorrelation structure and the skewness of processing times on the variance rate and the inter-departure variance. In particular, it was observed by simulation that by increasing the skewness of the processing times the inter-departure variance also increases. The complexity of Hendrick's approach is comparable to Miltenburg's [61], being dependent on the number of states representing the system.

Tan made a series of studies on the output variability of production systems. The works include the analysis and calculation of the output variability for machines in isolation, multi-stage unbuffered lines, production lines with parallel and series machines and small buffered manufacturing systems. He proposed both continuous time models [77, 79, 80] and discrete time models [81, 82, 84] for the analysis. In terms of investigated machine models, the studies include reliable machines with exponential processing times [77], unreliable machines with a single failure mode featuring geometrically or exponentially distributed failure times [82] and unreliable machines with Coxian distributed repair time [79]. Performance measures discussed include the asymptotic variance rate and the service level of the system.

Concerning the analysis of multi-stage systems, Tan proposed a matrix-geometric method for the estimation of the asymptotic variance rate in two machine lines with single failure mode machines and a finite buffer. Compared to Miltenburg's approach, the method proposed by Tan is much more efficient in terms of number of executed floating point operations. Tan uses the same approach for evaluating the variance of two-stage production lines with single failure mode machines as a function of time, again by exploiting the special structure of the transition matrix [81]. The method uses the Grassman approach [39] to iteratively obtain the performance of interest. The complexity of the adopted procedure depends both on the length of the observed time period and on the size of the Markov chain describing the process, thus on the buffer capacity. Tan [84] increases the computational efficiency of his algorithm allowing evaluating the performance of multi-stage production lines with unreliable machines and finite buffer capacity by using an exact procedure. Moreover, he studies the variability of the output for production lines controlled by different policies such as

Kanban, Basestock and CONWIP.

Ou and Gershwin [69] obtain closed form expressions of the variance of the lead time in a two machine line in which machines may fail in only one mode. Gershwin proposes a method for the calculation of the variance of the output of a single machine with a single failure mode in closed form Gershwin [28]. His method is based on the solution of the difference equation describing the system dynamics. The developed method is then used to approximate the performance of production lines through a decomposition approach. The effect of previous stages on the last machine in the system is considered by adjusting the failure and repair parameters of the single machine model. However, the method is shown to have large errors in the variance estimation (around 20% compared to simulation results) since the adopted decomposition equations [29] did not capture and propagate the output variability throughout the line. Carrascosa [11] extended the method of Gershwin to the case of the isolated machine with multiple failure modes.

Li and Meerkov [51] studied the variance of the output for production lines composed of unreliable machines and finite buffers. The most limiting assumption to the application of their method is the Bernoulli reliability model, which assumes repair time equal to the cycle time of the machines. Section 4.1.5 analyzes the features of such reliability model in depth. The authors focused on the "due time performance" which is an equivalent measure of the service level. Recently, more complex machine models providing insights about the transient behavior of the system have been studied in depth [52] and [60].

Other works that studied the transient behavior include [25] and [12]. In fact,the work of Dincer and Deler [25] studied both the transient and steady-state variability in the output of small buffered lines with reliable machines featuring exponentially distributed processing times, by adopting n-fold convolution of the inter-arrival and the processing time distributions. Chen and Yuan [12] focused on the system output mean and variance during the transient period. The approach models long unbuffered production lines with unreliable machines subject to a single failure mode, with exponentially distributed failure and repair times, by using a sample path method.

Ciprut et al. [14] used a fluid Markovian model to derive an exact closed-form formula to calculate the first two moments of the asymptotic output for unreliable machines with generally distributed up and down times. An attempt to extend this approach to two-machine one buffer dipoles was made, by approximating the dipole behavior with an equivalent single machine having two switching operational modes. When the second machine is not starved, the equivalent machine is exactly identical to the second machine of the dipole; when the second machine is starved, the equivalent machine behaves as the first machine in the dipole. However, the autocorrelation structure of the starvation times was not considered, thus this approximation may perform poorly in specific configurations. The exact analysis of Ciprut et al. [14] was recently extended by Angius et al. [4] to handle any system modeled as continuous and discrete time reward models, including machines with general Markovian structure.

Recently, other approximate methods for the analysis of the output variance in long multi-stage lines were introduced. He et al. [40] studied serial buffered multi-stage systems, with reliable machines featuring exponential processing times. The approximate method relied on the exact Markovian arrival process analysis of a simplified two-station one buffer sub-system and a compression method through propagate the output variance along subsystems. The difference with decomposition approaches is that the compression (also called aggregation) algorithm does not iterate backwards, form the last subsystem to the first. The system behavior is analyzed and it is shown that the variance of the output always increases with the buffer capacity, for this type of systems. However, no estimation of the method accuracy towards simulation is given in this paper. Another approximate method was proposed by Manitz and Tempelmeier [55], who studied long assembly lines with finite buffers and general service times. Their approach used a two-moment approximation to estimate the output variability in the assembly line, by measuring the coefficient of variation of inter-departure time.

Finally, the effect of the autocorrelation structure proposed by Hendricks [42] have been further investigated by Colledani et al. [17], for small systems featuring unreliable machines affected by multiple failure modes. They also managed to evaluate approximately the asymptotic variance rate of multi-stage production lines with machines having multiple geometric failure modes [19]. The proposed decomposition method suffers the same limitations of the decomposition method proposed in Gershwin [28]

Most of the aforementioned approaches were compared in Tan [82, and references therein] where a summary of papers that dealt with the output variability in production lines prior to year 2000 can be found. Tan's paper presents a classification of the existing methods based on the considered system layout, the considered system parameters and their distributions, the type of solution adopted, exact or approximate, and the complexity of the system that can be analyzed. According to this analysis, there are few methods that consider the problem of analyzing the output of multi-stage production lines with unreliable machines and finite buffer capacity. These available methods only consider the assumptions of exponentially or, in the discrete time domain, geometrically distributed machine failure and repair times. However, while in real systems the times to machine failures can often be modeled using exponential or geometrical distributions with acceptable accuracy, given the mechanical and electronic nature of failures, the times to repair are rarely observed to follow exponential distributions [45]. This thesis revises this critical assumption of the existing methods by considering general Markovian machines in the analysis.

2.4 Progress Beyond the State of Art

The proposed work fills the following gaps in the literature:

- *The exact evaluation of general Markovian small systems:* Methods that analytically calculate the output variability in a manufacturing system use simple machine structure models. This work proposes the General Markovian model in discrete time. This machine model is complex enough to model real failure modes happing at manufacturing lines.

- *The output variability evaluation of multi-stage production lines* : This work proposes methods that are able to calculate the output variability measures for long multi-stage manufacturing lines with an acceptable accuracy and speed.

- *Optimization:* The problem of optimization of second order moments of the output was not tackled previously. In this work, we will present a non-linear gradient method that is able to optimize the performance measures of the line.

- *Application to Industrial case:* A case study is analyzed explaining in detail the process for evaluating the output variability of a real production line.

Chapter 3

Production System Model

In this chapter, the characteristics of the production systems studied in this work are presented. Moreover, the terminology, notation, and the basic assumptions are detailed. The taxonomy presented in this chapter will be considered as a reference all over the thesis. This chapter has been summarized from the following references Gershwin [29], Papadopoulos et al. [70], Li and Meerkov [52], and Colledani [20]

3.1 Production Systems Architecture - Transfer Lines

The main focus of this thesis is directed to analyzing the output variability of transfer lines. In this section the definition and the adopted terminology for transfer lines will be introduced.

3.1.1 General Properties of Transfer lines

Synchronous and Asynchronous Lines: in synchronous production lines cycle times of different machines are identical and deterministic; therefore operations start and stop contemporarily for each machine. In asynchronous production lines cycle times may differamong machines and operations do not start and stop contemporarily for each machine.

Discrete and Continuous lines: in discrete production each operation requires a fixed time to process a part and the number of products present in buffers, at each time instant, is an integer number. Typical applications of discrete systems can be found in automotive lines, white goods production lines and mechanical components production lines. In continuous production systems machines perform operations on continuously flowing incoming parts. In this case, the quantity of products stored in buffers is a real number. Typical applications of this type of systems can be found in food industry, textile production lines, chemical lines and pharmaceutical lines. This are commonly analyzed through the use of continuous models, which treat the flow of material as a continuous fluid. These continuous models can also approximate the behavior of discrete systems in the case in which cycle times are consistently shorter than failure and repair times and buffers are small in size.

Starvation and Blocking phenomena: a machine is said to be starved if no part is available for processing. The machine is said to be blocked if there is no available place to store the processed part. Blocking and starvation phenomena are usually caused by interruptions of flow which propagates through the line. If no buffers are present, a failure of a machine immediately propagates to all the other machines composing the line. In this case, maximal coupling of machines in the line is verified. Buffers are commonly adopted in real production systems to decouple the behavior of machine and prevent blocking and starvation phenomena from propagating along the line. Once one machine fails, starvation propagates to the downstream machines while blocking propagates to the upstream machines. Therefore, machine M_i is blocked by a failed downstream machine M_j if all the buffers among M_i and M_j are full. On the other hand, M_i is starved by an upstream failed machine M_k if all the buffers between M_k and M_i are empty. Thus infinite buffers makes the propagation of blocking to upstream machines impossible, while still allowing downstream machines to be starved.

Capturing the correct dynamic of propagation of blocking and starvation in the system is fundamental for the development of accurate models and methods for the performance analysis of systems. Two models are generally adopted to describe the blocking and starvation dynamics [23]: they are known as Blocking Before Service (BBS), also named production blocking, and Blocking After Service (BAS) also named communication blocking. The first mechanism considers the following dynamics: the machine starts processing a product only if place to store it is available in the downstream buffer. The second mechanism models the case in which the machine starts processing the part (taking it from the upstream buffer) and, once finished the operation, the availability for place to store the part in the downstream buffer is checked. Analytical methods generally adopt the BBS assumption.

Saturated Systems: it is a system in which infinite supply of raw workpieces is available upstream the first machine in the line and an unlimited storage area is available downstream the last machine in the line. In these systems, the first machine is never starved and the last machine is never blocked.

Open and Closed Systems: in open systems the arrival of parts at the first machine and the departure of products at the last machine are independent events. In closed systems, as soon as a finished part is released by the system, a raw workpiece starts its processing on the first machine. If a system is closed it is populated by a constant amount of circulating products.

Linear and Non-Linear flow systems: if the flow of parts throughout the system comes from one source i.e. as in transfer lines, the system is considered linear. Otherwise, if the system has actuators that perform split/merge operations i.e. as in parallel lines, the systems is called Non-linear.

Assembly/ Disassembly systems: lines in which some of the manufacturing machines perform assembly and disassembly operations. The system presents a fork and join structure, more than one buffer can be located downstream the same machine.

3.1.2 Transfer Lines

The focus of this work is directed to saturated, linear and open transfer lines modeled in discrete time; parts arrive to the first station and move after being processes to the following one, then depart the system. In addition, the flow within the production line is linear, i.e. no split/merge operation and no assembly/disassembly machines are used. Finally, quality issues i.e imperfect production, quality failures, out of control and rework are not considered.

Material transfer within the system is represented by arrows. Since transfer lines are push system i.e. machines process available existing parts in the upstream buffer space as long as they are operating and not blocked (BBS is assumed). Buffers are represented by circles. Figure 3.1 shows a representation of a machine connected to two buffers.

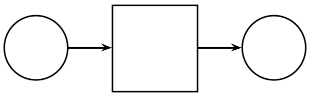


Figure 3.1: An example of a machine connected with two buffers

3.1.3 The Structure of a Transfer Line - Assumptions used

In this work, any reference to production system or manufacturing system refers to transfer lines. The production system has a simple and linear structure as seen in Figure 3.2. It is generally composed by K machines and $K-1$ buffers. Machines are denoted with M_k with $k = 1 \dots K$ $(M_1, M_2, M_3, ...M_K)$. The buffer placed downstream of a machine M_k is denoted as B_k, with $k = 1, \dots K-1$. Raw parts enter the system upstream the first machine M_1, then are processed through all the K stages and leave the system downstream the last machine M_k.

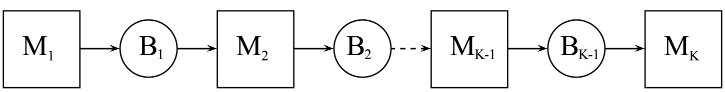


Figure 3.2: Representation of transfer line

Application of this type of production systems are found in several fields and sectors. An example from the automotive sector is the engine block machining line analyzed in Scania [18]. The line is consisting of 22 machines and 21 buffer. The architecture has a **S** shape. The system is reported in Figure 3.3.

Another example is from the mineral water bottling industry, the manufacturing plant of Levissima 0.5 Liter bottles, and it is reported in Chapter 8. The line consists of 7 machines and 6 buffers. The

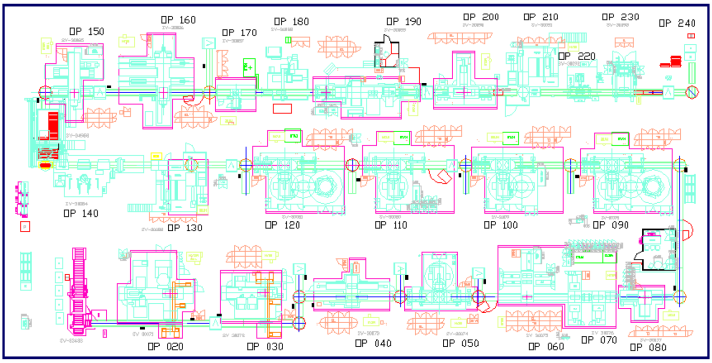


Figure 3.3: Real transfer line producing the engine block at Scania

architecture has a complex shape. The system is reported in Figure 3.4.

3.2 Machine Characteristics

This Section provides the taxonomy of the main characteristics of the different machine types which compose the considered set of systems. With this purpose, modeling features of operational failures are provided. Furthermore, Table 3.1 shows a list of notations that will be used throughout this work.

3.2.1 General Properties of Machines

Production in Isolation Mode: a machine is said to produce parts in isolation mode if it is considered to be isolated from the other machines or buffers in the line, i.e. if it is not affected in its behavior by blocking and starvation phenomena, as if it is connected to infinite raw material buffer upstream and infinite space buffer downstream.

Cycle time: the time required for a machine to process a product, while working in isolation, is named cycle time. It can be deterministic, if it is not varying from one part to the next, concerning a given process. It is stochastic, if it is randomly varying from one part to the next.

Operational Failures: operational failures are those disturbances which cause the immediate interruption of the production flow for a machine.

Table 3.1: Notations used throughout the thesis

System Parameters	
K	Number of machine in the line
i	Machine index
M_i	i^{th} machine in the line
B_i	i^{th} buffer in the line
N_i	Capacity of buffer B_i
MTTF_{ij}	Mean Time to Failure of failure mode j in machine M_i
MTTR_{ij}	Mean Time to Repair of failure mode j in machine M_i
e_i	Efficiency in isolation of machine M_i
v_i	Asymptotic output variance rate of machine i.
$\boldsymbol{\lambda}_i$	Transition probability matrix for machine i
$\boldsymbol{\mu}_i$	Binary reward output vector for machine i
E	Efficiency in isolation of the system.
V	Asymptotic output variance rate of the system.

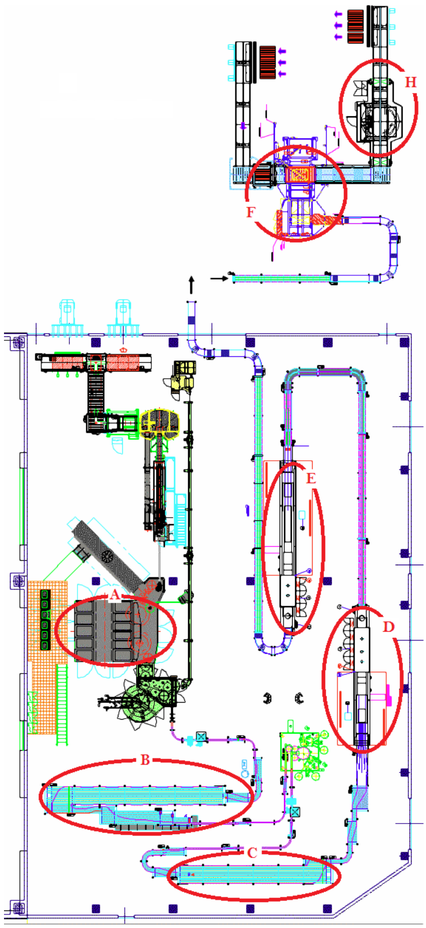


Figure 3.4: The layout of the production line of Levissima

Failures which stop the whole production of the system, like energy provision interruptions, are not considered among these types of failures, thus the independence of failures among different machines is considered. In order to restore the machine to the operative conditions, the intervention of an operator is required. Two types of failures are generally observed in real production systems, i.e. Operation Dependent Failures (ODF) and Time Dependent Failures (TDF).

- **Operation Dependent Failures (ODF):** are those failures that can happen only if the machine is operational, i.e. not starved nor blocked. This are typically mechanical failures, such that the tool breakage, the errors of sensors while positioning the work piece in the work zone, the lack of material and mechanical jamming.

- **Time Dependent Failures (TDF):** are those failures that can happen even if the machine is starved or blocked, i.e. the failure occurrence do not depend on the machine state. They are typically electronic failures, such that light burn-outs, machine screen problems and machine communication problems.

Failure models: machines operating in real production systems are characterized by different modes of failure. The most complex structure analyzed in this work is the general Markovian machine model with a binary reward. Other models include the multiple failure mode geometric machines, single failure mode geometric machine and the Bernoulli machine, explanation and the performance evaluation of such models is presented in Chapter 4.

3.2.2 Buffer Characteristics

Buffers are present in real production systems with the role of transporting material from one machine to another one, decoupling the behavior of the machines and reduce the effect of the propagation of blocking and starvation in the line. They can be automatic conveyors, floor space, ...etc. This work considers buffers to be reliable. and not subject to failure.

3.3 General System Modeling Assumptions

In this section, the list of assumptions and notation common to the analytical methods proposed in this thesis is provided:

System architecture assumptions:

- Saturated and Open systems are considered.
- Discrete systems are considered. Stations are characterized by a deterministic processing time, scaled to one time unit. The flow of material is considered as discrete.
- Blocking before service is considered.

- Perfect production, quality issues not considered.
- The following convention is adopted: in the same time unit, machine state transitions (like failures and repairs) occur at the beginning of the time unit and buffer level is updated at the end of the time unit. See Figure 3.5.

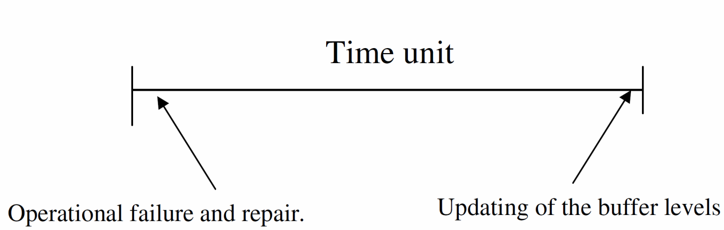


Figure 3.5: Adopted convention on failures and buffer level updating

Machine assumptions:

- Unreliable, general Markovian systems are considered.
- Machines' output is characterized by a binary reward vector μ, indicating which states of the machine are up and which are down.
- Failures are Operational Dependent Failures (ODF).
- The machine can have down states, up states, and transitions among all possible states.
- Transition from state j_1 to j_2 occurs with probability P_{j_1,j_2}, and follows the geometric distribution with mean $({}^1/P_{j_1,j_2})$, these transitions are ordered in the transition probability named λ.
- machines produce perfect parts.

Buffer assumptions:

- Buffers have finite capacity N.
- Transient time is zero.
- Buffers are perfectly reliable.

Chapter 4

The Analysis of the Output Variability of Small Production Lines

4.1 The Analysis of the Output Variability of Single Machines in Isolation

4.1.1 Introduction

In this chapter, the calculation of different output variability measures for small Markovian system will be presented. The Markovian system structure is very useful for the analysis of complex machines and small lines. In the literature most of the models presented in discrete time used geometric single failure mode machines. In reality machines can have more complex failure and repair dynamics, thus can not be modeled with the geometric machine. Hence, the general Markovian system can be used in such cases. This chapter will present the building block of the analysis presented in following structures that depends on the decomposition approach.

This chapter will present a method for the calculation of the first two moments of the general Markovian system. Then this system model can be utilized to model some specific cases of machine structures studied previously, namely, machines with Ph-Erlang down machines, machines with Coxian down machines, the multiple failure modes geometric machine [19], the single failure mode geometric machine [28, 84], the Bernoulli machine [52].

4.1.2 General Markovian Systems

In this section, a discrete time system with s different states and an underlying transition probability matrix $\mathbf{P}$ is considered. The system has a constant processing cycle time. Time is scaled so that the processing cycle time is one time unit. The system is characterized by down states, up states and transitions among all possible states. Transitions from state j_1 to j_2 occurring with probability (p_{j_1,j_2}) follow the geometric distribution with a mean $(1/p_{j_1,j_2})$. By convention, transitions can happen only at the beginning of a time unit. The system has a binary reward column vector $\boldsymbol{\mu}_{s\times 1}$ that governs its output. The reward μ_j assumes the value *one* if j is a productive state, and *zero* otherwise, with $j = 1, \ldots, s$.

Description of the method

The machine output is a binary random variable Y_i taking the value *one* if the machine produces a piece in period i and *zero* otherwise. The mean production rate e for this machine is calculated as follows:

$$e = \sum_{j=1}^{s} \pi_j \mu_j \tag{4.1}$$

where $\pi(j)$ is the steady state probability of being in state j.

The total production output Z_t of the machine at time t is given by the sum of the its output:

$$Z_t = \sum_{i=1}^{t} Y_i \tag{4.2}$$

and $var[Z_t]$ is calculated by:

$$var\left[Z_t\right] = \sum_{i=1}^{t} var\left[Y_i\right] + 2\sum_{i=1}^{t}\sum_{l=i+1}^{t} cov\left[Y_i, Y_l\right] \tag{4.3}$$

which is the sum of two different components [42]. The first component is related to the variance of the single random variables Y_i, while the second component arises when the series Y_i are not independent but timely autocorrelated. Since we are interested in calculating the steady state performance, we assume that the output process is stationary at the beginning of the analyzed time interval. Thus, equation (4.3) can be rewritten as:

$$var\,[Z_t] = t\sigma_Y^2 + 2\sum_{k=1}^{t-1}(t-k)cov_k\,[Y] \tag{4.4}$$

where Y is the random variable of the stationary output process, and $cov_k[Y]$ is the autocovariance of lag k of the time series Y.

In order to apply equation (4.4), it is necessary to know the variance and autocovariances of the process output in steady state, i.e. to calculate σ_Y^2 and $cov_k[Y]$. By definition, the variance is:

$$\sigma_Y^2 = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = e - e^2 \tag{4.5}$$

because Y is binary, thus $\mathbb{E}[Y^2]$ is equal to $\mathbb{E}[Y]$. The interpretation of σ_Y^2 is straightforward, if the machine is observed in steady state n times independently, the variance of the observed Y values tends to σ_Y^2 as $n \to \infty$. By definition, the autocovariance of lag k is:

$$cov_k[Y] = \mathbb{E}[Y_i Y_{i+k}] - \mathbb{E}[Y_i]\mathbb{E}[Y_{i+k}] \tag{4.6}$$

Again since Y is binary, $\mathbb{E}[Y_i Y_{i+k}]$ reduces to the probability that the machine is up both at periods i and $i+k$. Therefore, the expression of the autocovariance of lag k becomes, after some manipulations:

$$cov_k[Y] = \sum_{j=1}^{s}\sum_{g=1}^{s}\pi_j\mu_j\mathbf{P}_{j,g}^k\mu_g - e^2 \tag{4.7}$$

The process tails off, i.e. as the lag increases the autocovariance approaches to zero for large values of k [76].

The spectral decomposition of the Perron-Frobenius theorem can be used to formulate the $\mathbf{P}^k$ matrix by means of the eigenvalues and eigenvectors of $\mathbf{P}$:

$$\mathbf{P}^k = \sum_{j=1}^{s}\lambda_j^k\eta_j u_j' \tag{4.8}$$

where $\lambda_1, ..., \lambda_s$ are the s distinct eigenvalues of $\mathbf{P}$, $u_1, ..., u_s$ and $\eta_1, ..., \eta_s$ the associated sequences of left and right eigenvectors respectively such that $u_r'\eta_j = 0$ if $r \neq j$ and $u_r'\eta_j = 1$ for all $r, j = 1, ..., s$[8].

By definition, the autocorrelation function of lag k is:

$$\rho_k[Y] = \frac{cov_k[Y]}{\sigma_Y^2} \tag{4.9}$$

Substituting equations (4.5), (4.7) and (4.9) into (4.4) and after some manipulations:

$$var[Z_t] = \sigma_Y^2 \left[t + 2\sum_{k=1}^{t-1}(t-k)\rho_k[Y] \right] \tag{4.10}$$

which is an exact formula for calculating the total output variance in a time period $[1, t]$. We can identify a number of k^* significant lags after which ρ approaches *zero* and the equation simplifies into:

$$var[Z_t] = t\sigma_Y^2 + 2\sigma_Y^2 \sum_{k=1}^{k^*}(t-k)\rho_k \tag{4.11}$$

The number of significant lags k^* is small compared to t, and it directly depends on the second largest eigenvalue (λ_2) of the matrix $\mathbf{P}$. The number of significant lags k^* that guarantee ρ_{k^*+1} is *zero* is calculated by:

$$k^* = \frac{\log \epsilon}{\log \lambda_2} \tag{4.12}$$

where ϵ is the required tolerance in the calculation of v [76]. The number of k^* lags increases as the size of $\mathbf{P}$ increases, indicating that the more complex the system is, the more states it has and the more time it takes to reach the steady state. The second largest eigenvalue also affects the transient system behavior [60]. Therefore, minimizing λ_2 corresponds to minimizing both the system warm up and v.

Equation(4.11) can be rewritten as:

$$var[Z_t] = t\sigma_Y^2 (1 + 2\rho_{tot}) - 2\sigma_Y^2 \sum_{k=1}^{k^*}(k \cdot \rho_k) \tag{4.13}$$

where $\rho_{tot} = \sum_{k=1}^{k^*} \rho_k$ is the total autocorrelation.
The other performance indicators can be calculated as:

$$v = \sigma_Y^2(1 + 2\rho_{tot}) \tag{4.14}$$

$$cv(t) \approx \sqrt{\frac{(1-e)}{e \cdot t}(1+2\rho_{tot})} \tag{4.15}$$

$$d = (1-e)(1+2\rho_{tot}) \tag{4.16}$$

This result is general and is not limited to only a unique failure mode or a single machine. It is also in accordance with the results of [42] for a simplified machine model.

The analysis of the output transient variance

The transient output variance is the variability observed in the output during warm-up period i.e the period needed to reach steady state [65, 60]. In the proposed approach the transient time can also be captured by equation (4.13). The advantage of this equation is that it divides the output variance into two terms the asymptotic output variance rate v and the transient component.

The analysis of the transient period is important for two reasons: the first reason is to know what is the time needed to reach the steady state time, while the second reason is how much does the output variance rate change during this period. The works of [60, 51] studied the transient period impact and length depending on the 2nd largest eigenvalue λ_2 only. In the proposed approach λ_2 is still used to determine the length of the transient period, however ρ_{tot} considers the impact of all the eigenvalues of the system and not only λ_2.

To study the effect of the transient period on the output variability, a machine with 60 states has been generated randomly with parameters $e = 0.705$ and $v = 2.4$. Figure (4.1) shows the changes in $v(i)$, with time i, where $i \ll t$ compared to the machine's v, it is interesting to notice that the value of v is higher than that of the transient behavior output variability. This indicates that considering v during design phase is more critical as it is always higher than $v(i)$ with the transient output variability component. This observation is valid for low transition probabilities that guarantee a positive ρ_{tot}, in fact, real machines abide to this rule.

In the following part of this chapter, the analysis of the output variability measures for some specific cases will be presented, namely, the multiple failure modes geometric machine, the single failure mode geometric machine, the Bernoulli machine.

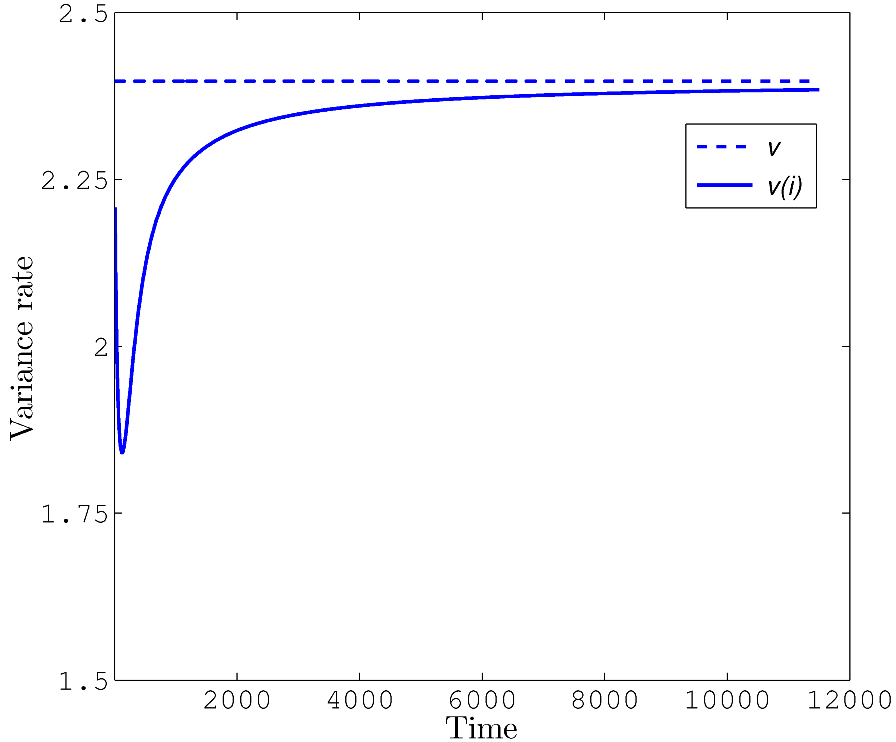


Figure 4.1: A comparison between the $v(i)$ and v for the studied 60 state machine

4.1.3 Isolated multiple failure modes geometric machine

The method proposed can be used for evaluating the output variability of an isolated multiple failure modes geometric machine [19]. In this case, the machine can be up (operational) in one mode or down (failed) in different modes $1, \dots f$ with failure and repair probabilities equal to $p_1, \dots p_f$ and $r_1, \dots r_f$, respectively. State transitions can happen between the up state and a down state only. Figure 4.2 presents the Markov chain of the multiple failure mode machines and the transition probability matrix $\mathbf{P}$ of this machine is:

$$\mathbf{P} = \begin{bmatrix} 1 - \sum_{j=1}^{f} p_j & p_1 & \dots & p_f \\ r_1 & 1 - r_1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ r_f & 0 & \dots & 1 - r_f \end{bmatrix} \tag{4.17}$$

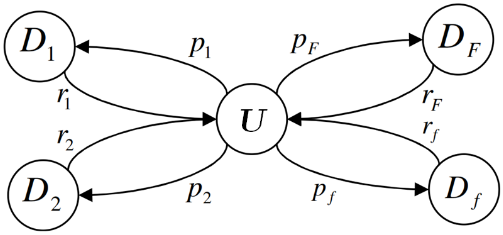


Figure 4.2: Markov chain representing the behavior of the multiple failure mode machine

The machine's mean production rate is [87]:

$$e = \frac{1}{1 + \sum_{j=1}^{f} \frac{p_j}{r_j}} \tag{4.18}$$

Equations (4.5) and (4.7) hold also for this case, with the only difference that e is calculated as in equation (4.18). Thus, after having calculated the variance and autocovariances of the time series Y, it is possible to calculate the variance of the cumulated production by using equation (4.11), while the other output variability measures can be calculated using equations (4.14) - (4.16).

4.1.4 Isolated single failure mode geometric machine

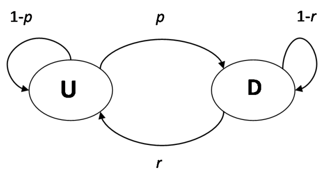
A sub-case of the previous example that has been widely analyzed in the literature is the case of isolated machine with single failure mode. In the geometric repair time case, this machine can be either up (operational) or down (failed) in a single mode as depicted in Figure 4.3(a) [28]. While operational the machine can fail with probability p at the beginning of the time unit. While failed, it can be restored with probability r, the matrix representing such system is depicted in Figure 4.3(b).

The expected value of the machine throughput is calculated as follows:

$$e = \frac{r}{p + r} \tag{4.19}$$

The variance of the efficiency is:

$$\sigma_Y^2 = e(1 - e) \tag{4.20}$$



(a) Markov chain for one machine

$$\mathbf{P} = \begin{bmatrix} 1-p & p \\ r & 1-r \end{bmatrix}$$

(b) Markov chain Matrix for one machine

Figure 4.3: Representation of a single machine with Markov chain

The coefficient of variation is:

$$cv_Y = \sqrt{\frac{1-e}{e}} \tag{4.21}$$

and, finally, the index of dispersion is calculated as:

$$d_Y = 1 - e \tag{4.22}$$

All these indicators refer to the random variable Y in the steady state. Furthermore, it can be noticed that all the variability indicators depend on e and not on the combination of p and r of the single machine. Figure 4.4 shows the value of these indicators as a function of e. The figure also shows that σ_Y^2 is always included between 0 and 0.25, while d_Y is a linear function starting from 1 and decreasing to 0, while cv_Y tends to $+\infty$ when $e \to 0$ and goes to 0 when $e \to 1$.

A well-known solution of this model is the steady state vector probabilities $\boldsymbol{\pi} = [\pi_{up}, \pi_{down}] = [e, 1-e]$. The expected cumulated production is:

$$\mathbb{E}[Z_t] = \sum_{t=1}^{T} \mathbb{E}[Y_t] = t\frac{r}{r+p} \tag{4.23}$$

The autocovariance has a special form, equation (4.7) becomes:

$$cov_k[Y] = e(1-e)(1-p-r)^k \tag{4.24}$$

In this case it is easy to see that the process tails off with an increasing k. Equation (4.8) becomes:

$$\mathbf{P}^k = \frac{1}{r+p}\begin{bmatrix} r & p \\ r & p \end{bmatrix} + \frac{(1-r-p)^k}{r+p}\begin{bmatrix} p & -p \\ -r & r \end{bmatrix}$$

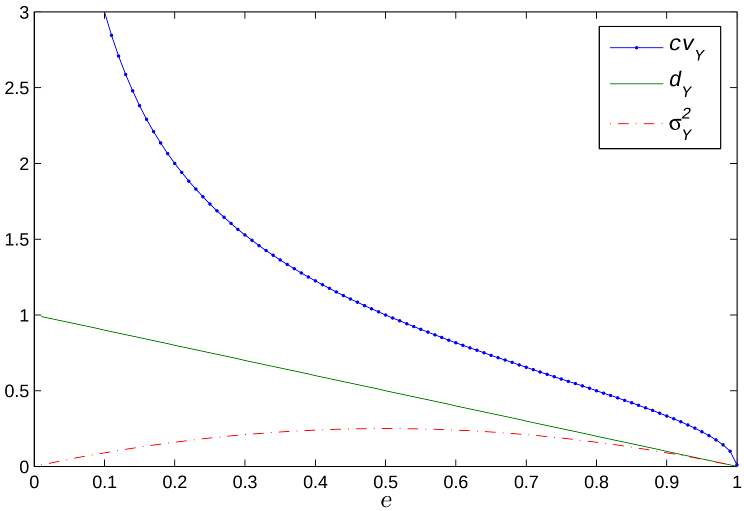


Figure 4.4: cv_Y,σ_Y^2 and d_Y as a function of e.

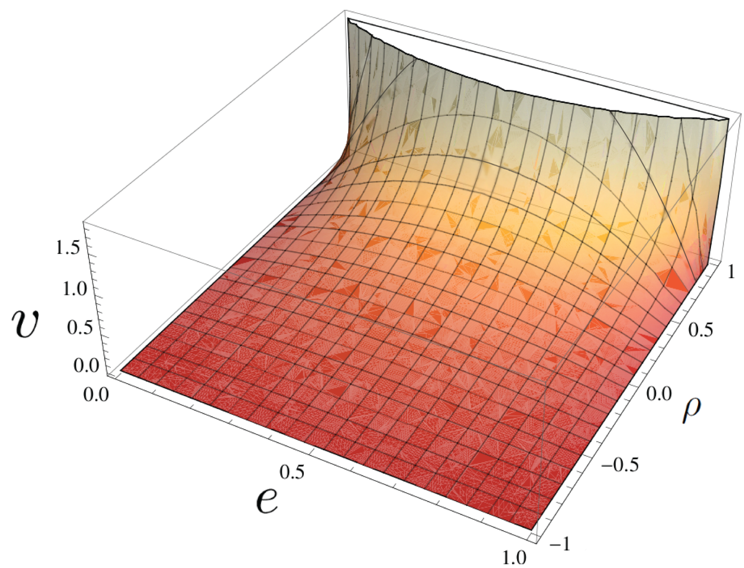


Figure 4.5: v vs e and ρ for the single failure geometric machine

where $\lambda_2 = 1 - r - p$ is the second largest eigenvalue of $\mathbf{P}$. In this case only, the autocorrelation function of lag k coincides with λ_2^k, and it is equal to:

$$\rho_k[Y] = \frac{cov_k[Y]}{\sigma_Y^2} = (1 - p - r)^k = \rho^k \tag{4.25}$$

substituting equations (4.5), (4.24) and (4.25) into (4.4) and after some manipulations:

$$var[Z_t] = e(1-e)\left[t + 2\sum_{k=1}^{t-1}(t-k)\rho^k\right] = e(1-e)\left[\frac{t - t\rho^2 - 2\rho + 2\rho^{t+1}}{(1-\rho)^2}\right] \tag{4.26}$$

which can also be written as a function of the single machine's variance as:

$$var[Z_t] = \sigma_Y^2 \cdot \left[\frac{t - t\rho^2 - 2\rho + 2\rho^{t+1}}{(1-\rho)^2}\right] \tag{4.27}$$

Finally, it is possible to calculate the other output variability measures as a function of e and ρ:

$$v = e(1-e)\frac{1+\rho}{1-\rho} \tag{4.28}$$

$$cv = \sqrt{\frac{1-e}{et}\left(\frac{1+\rho}{1-\rho}\right)} \tag{4.29}$$

$$d = (1-e)\frac{1+\rho}{1-\rho} \tag{4.30}$$

Impact of machine reliability parameters

The proposed method is used to derive insights on the behavior of the output variability under changes in the main system parameters. Firstly, the analysis of the impact of the machine reliability parameters is carried out. Figure 4.5 shows the behavior of v for the single failure mode geometric machine, with different values of e and ρ. It can be noticed that ρ impacts v more than e does when ρ is higher than 0. Figure 4.6 shows the relationship between v and e as a function of p and r. The behavior of v depends on relative position of the machine respect to r^* and p^* curves which divide the plot into three regions A, B and C. The curves r^* and p^* are calculated by:

$$r^* = 2 - \sqrt{4 + p^2 - 2p} \tag{4.31}$$

$$p^* = 1 - Re\left(\sqrt{1 - 4r + r^2}\right) \tag{4.32}$$

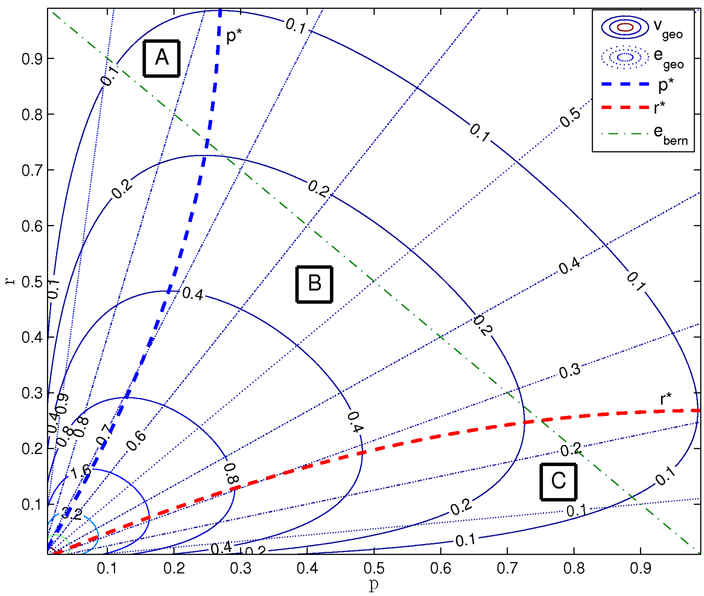


Figure 4.6: Throughput and asymptotic variance rate as a function of p and r for the geometric machine and for the Bernoulli machine (as a function of p).

where $Re(\cdot)$ is the real part of the function. The behavior of changing p and r in each region has the following effect:

- *Region A*: $p \leq p^*, \forall r$. Increasing r or decreasing p has a double positive effect, i.e. v decreases and e increases. In other words, actions that increase machine MTTF or that decrease the MTTR have a positive effect on both e and v.
- *Region B*: $p \geq p^*, r \geq r^*$. Increasing r has the same double positive effect on v and e, while decreasing p has the positive effect of increasing e, coupled with the negative effect of increasing v. In this region a rather counterintuitive effect is observed. If the MTTF related of the machine is increased, for example by applying machine improvement plan, the throughput of the system is increased but the output becomes more unstable since the output variance is higher. On the contrary, up to a certain extend, keeping the machine down for a longer time may be beneficial since the loss in the throughput is compensated by a higher stability of the output.
- *Region C*: $\forall p, r \leq r^*$. Increasing r or decreasing p has the positive effect of increasing e, coupled with the negative effect of increasing v. This is a trade-off region where the only possibility of decreasing the output variability of the machine is to decrease its throughput. Therefore, up to a certain extend, keeping the machine up for a shorter time may be beneficial since the loss in the throughput is compensated by a higher stability of the output. It should be noticed that machines in this region can drastically affect the system performance. Thus, it would be beneficial to implement improvement actions that upgrade the machine to regions A or B.

Therefore, this map can be used to select proper machine reconfiguration actions that improve the performance both in terms of asymptotic throughput and variance rate, depending on the position of the machine in the graph. Knowing the relative position of a machine from the curves r^* and p^* would help machine designers and production managers identifying the best improvements actions for a machine.

Impact of the repair time distribution

The impact of the repair time distribution on v and e is investigated next. The studied machines have a geometric time to failure with mean equal to 10 time units. Three different repair time distributions were

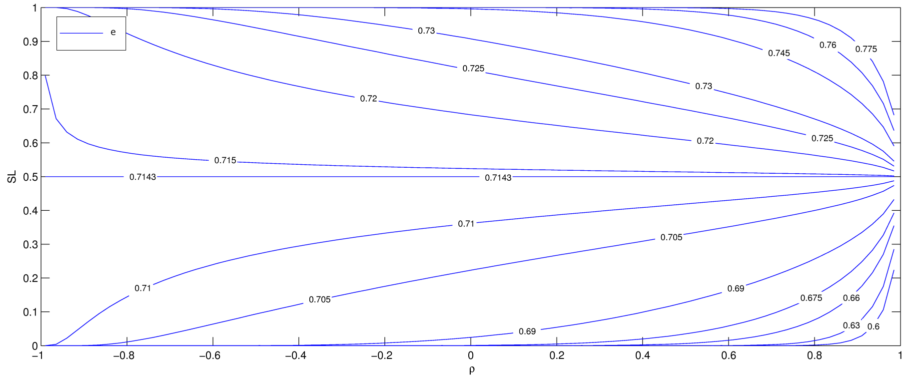


Figure 4.7: Service level as a function of ρ and e when the demand x is equal to $10,000$ products within a period of $14,000$ cycle times

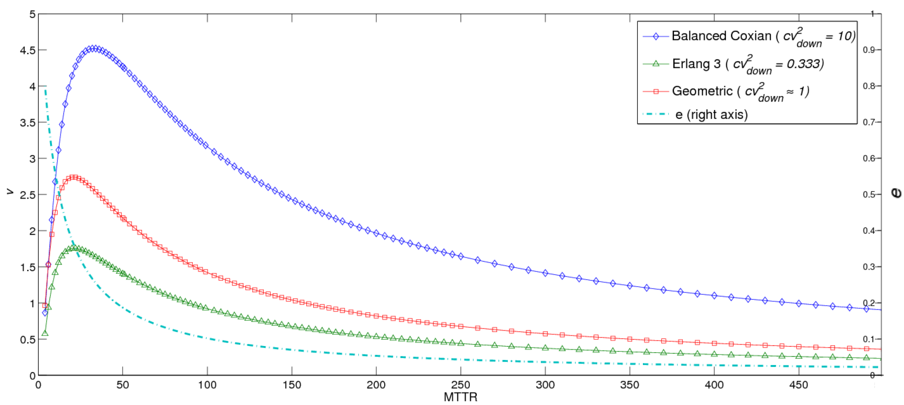


Figure 4.8: v and e as functions of MTTR and MTTF for different machine reliability models

tested with the same mean but different square coefficient of variations (cv^2_{down}), i.e. the geometric distribution ($cv^2_{down} = 1$), the balanced mean Cox-2 distribution ($cv^2_{down} = 10$) and the Erlang distribution ($cv^2_{down} = 0.333$). Then the MTTR is varied in the range $[1 - 500]$. In order to calculate the asymptotic variance rate of non-geometrical machines, the general method explained in Section 4.1.2 has been adopted by using the state-space representation of the phase-type distributions mentioned above.

The results are shown in Figure 4.8. It can be seen that the distribution of the time to repair affects the output variability of the machine but not the average performance. More interesting, the v is a concave function showing a maximum when the reciprocal of MTTR crosses the r^* line. It can also be noticed that the cv^2 of the TTR affects the position of the curve r^*. Specifically, while the cv^2_{down} increases, the maximum is visible for lower MTTRs. Furthermore, higher cv^2_{down} entails higher production variability. Therefore, in order to understand the behavior of the output variability of the system as a function of MTTR a second moment analysis should be performed.

Impact of the autocorrelation coefficient on the service level

Figure 4.7 shows the impact of ρ on the service level for certain values of e and a fixed demand x equal to $0.7143 \cdot t$. When e is equal to $^x/_t$, the SL is always 50% regardless of the amount of correlation in the output. When e is greater than $^x/_t$ increasing ρ will cause SL to decrease, as the probability of obtaining a long consecutive series of no output (i.e. $Y_i = 0$) increases. While when e is smaller than $^x/_t$, increasing ρ will cause SL to increase, as the probability of having long series of consecutive outputs (i.e. $Y_i = 1$) is higher.

Matching the first two asymptotic moments of the output by a geometric machine

The output variability analysis of subsection 4.1.4 can be reversely used to match the first two model of the output of a complex manufacturing system with a geometric single failure machine model. Let's assume that a complex manufacturing system produces parts with a characteristic throughput e and a cumulative autocorrelation coefficient of the output process ρ_{tot}. The parameters p and r of the geometric single failure mode machine matching the same first two asymptotic moments of the output can be obtained with the following equations:

$$\begin{cases} p_{eq} = \frac{1-e}{\rho_{tot}} \\ r_{eq} = \frac{e}{\rho_{tot}} \end{cases} \quad (4.33)$$

Equation (4.33) can be used to find an equivalent geometric machine on the basis of estimates for e and ρ from real field data. Moreover, it could be used for propagating both the first and second asymptotic moments of the output between the different subsystems within a new decomposition technique to analyze long production lines [see [64]]. This extension will be subject of future research activities.

The Effect of machine cycle time on the output variability

To study the effect of the cycle time on the output variability, we can consider one of the machines presented in Scania case of Chapter 1. For example machine 130, which is located in the middle of the line, had a MTTR and MTTF equal to 33.22 min and 68.47 min, respectively. The machine is modeled as a single failure geometric machine. The cycle time of this machine was 4.13 min, while its efficiency in isolation was 0.6734. Figure 4.9 and Table 4.1 show the effect of changing the cycle time on the output asymptotic variance rate v for machine 130. It can be seen that although machines with the same values of MTTF and MTTR have the same efficiency in isolation, they produce different output variability depending on their cycle time.

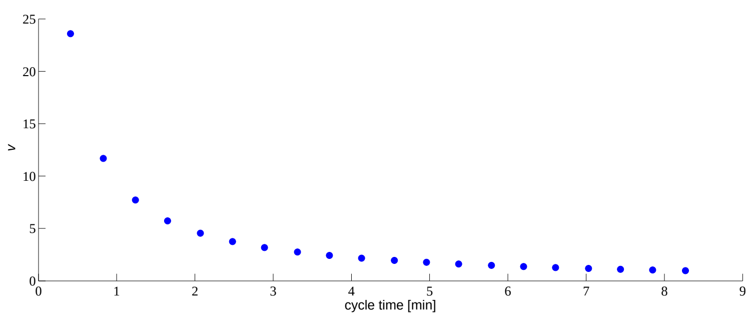


Figure 4.9: The effect of changing the cycle time on the asymptotic variance rate for a single machine

Table 4.1: The effect of changing the cycle time for a machine from the Scania case with MTTR and MTTF equal to 33.22 min and 68.47 min, respectively

Cycle Time(min)	p	r	e	v
0.4133	0.0060	0.0124	0.6733	23.5852
0.8266	0.0120	0.0248	0.6733	11.6826
1.2400	0.0181	0.0373	0.6733	7.7151
1.6533	0.0241	0.0497	0.6733	5.7313
2.0666	0.0301	0.0622	0.6733	4.5410
2.4800	0.0362	0.0746	0.6733	3.7475
2.8933	0.0422	0.0871	0.6733	3.1807
3.3066	0.0482	0.0995	0.6733	2.7556
3.7200	0.0543	0.1119	0.6733	2.4250
4.1333	0.0603	0.1244	0.6733	2.1605
4.5466	0.0663	0.1368	0.6733	1.9441
4.9600	0.0724	0.1493	0.6733	1.7638
5.3733	0.0784	0.1617	0.6733	1.6112
5.7866	0.0845	0.1742	0.6733	1.4804
6.2000	0.0905	0.1866	0.6733	1.3670
6.6133	0.0965	0.1990	0.6733	1.2678
7.0266	0.1026	0.2115	0.6733	1.1803
7.4399	0.1086	0.2239	0.6733	1.1025
7.8533	0.1146	0.2364	0.6733	1.0329
8.2666	0.1207	0.2488	0.6733	0.9703

4.1.5 Bernoulli machine

The single failure mode geometric machine can be reduced to the Bernoulli machine when $p + r = 1$ [51]. Assuming the random variable Y is in the steady state. The mean production rate, denoted e, of the machine under the Bernoulli model is given by:

$$e_Y = 1 - p \tag{4.34}$$

The variability performance measures for the efficiency of the Bernoulli machine are:

$$\sigma_Y^2 = e(1-e) = (1-p)p \tag{4.35}$$

$$cv_Y = \sqrt{\frac{1-e}{e}} = \sqrt{\frac{p}{1-p}} \tag{4.36}$$

$$d_Y = (1-r) = p \tag{4.37}$$

It could be noticed that all the output variability measures depend on p alone, thus they have a very simple closed form. Equations from 4.34 to 4.37 can be written in terms of r substituting p with $1 - r$ too.

The autocorrelation among the rewards of the Bernoulli machine does not exist because the random variables Y_k are independent. Therefore, the performance indicators related to the cumulated production can be calculated by setting $\rho = 1 - p - r = 0$. Figure 4.6 shows the linear behavior of e vs p, while in figure 4.5 we can see that when ρ is equal to 0, v is equal to $e(1-e)$. The expected valued of Z_t is simply calculated as:

$$\mathbb{E}[Z_t] = t \cdot e = t \cdot r = t(1-p) \tag{4.38}$$

and the $var[Z_t]$ becomes:

$$var[Z_t] = r \cdot p \cdot t = (1-p)p \cdot t \tag{4.39}$$

the other output variability measures become:

$$v = r \cdot p = p(1-p) \tag{4.40}$$

$$cv = \sqrt{\frac{p}{t \cdot r}} = \sqrt{\frac{p}{t \cdot (1-p)}} \tag{4.41}$$

$$d = (1-r) = p \tag{4.42}$$

4.1.6 Analysis of the speed of the method

In order to examine the speed of the proposed method, different randomly generated isolated machines with an increasing number of states (up to 100 states) were evaluated analytically. In the first part, k^* was calculated from equation (4.12) then the algorithm calculates all the ρ's describing the evolution of the system and uses equation (4.14) to calculate v. Table 4.2 shows the output measures and the evaluation time of different machines, the evaluation time for a machine with 100 states the is only around 3 minutes. Furthermore, the method depends on the length of the transient period of the system, if the machine has long transient period i.e. k^* is high the evaluation time is high too, and vice versa.

Table 4.2: The speed of the proposed method

Number of States	e	v	k^*	Evaluation Time
2	0.9319	0.1092	14	0.143
3	0.8507	1.3500	110	0.011
5	0.8524	0.9746	79	0.031
10	0.8200	0.8358	342	0.086
15	0.7959	1.1075	523	0.129
20	0.7330	4.3112	875	0.313
25	0.7426	3.3546	1322	0.767
30	0.8143	2.9628	1169	0.966
40	0.7140	13.3171	1641	2.141
50	0.7064	3.4715	2141	4.601
60	0.7215	1.27629	6259	19.230
70	0.7308	12.7521	10064	41.361
100	0.7012	16.9931	22538	193.688

4.2 Two Machine Lines

The system is composed of machines and an intermediate buffer with limited capacity. The approach proposed in Subsection 4.1.2 applies also in this case. The only difference is that the Markov chain underlying this system is more complex in the number of states, being dependent on the number of up and down modes of both machines and on the buffer capacity. In particular, the system state is identified by the triple $\boldsymbol{x} = (n, \alpha, \beta)$, where n indicates the buffer level, α and β are the states of the first and second machine, respectively.

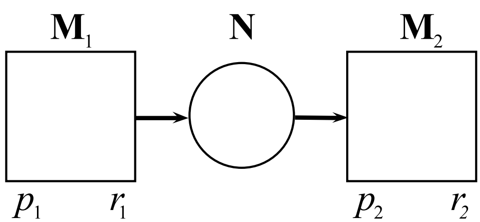


Figure 4.10: The studied single failure mode two machine line

Focusing the attention on the second machine of the tandem line, it is assumed that the random variable Y_i is equal to 1 if the observed machine produces one piece in period i, and 0 otherwise. In particular Y_i can be null for two different reasons: the second machine is down, or it is starved because the first machine is down and the buffer is empty. The expected throughput of this system is the sum of all the steady state probabilities in which the second machine is operational and not starved [29]. Therefore, the variance is calculated using equation (4.5) and the autocovariance of lag k is calculated using equation (4.7).

Again, the system is assumed to be stationary. Thus again equation (4.4) can be used for the calculation of the variance of the cumulated production. The time needed to evaluate such system depends on the total number of states and the significant number of autocorrelation lags k^*. Using sparse representation of matrix $\mathbf{P}$, the methods depends on the multiplication of the sparse transition matrix of size $(N+1)(s_1)(s_2)$ for k^* times, where s_i is the number of states of machine i (with $i = 1, 2$).

To be consistent with the notation adopted in the available literature in the field, we will use e and v for isolated machines and E and V for multi-stage systems.

The behavior of V in a two machine system is complex. Carrascosa [11] has shown that the shape of V curve as a function of buffer capacity (N) is very sensitive to the machine parameters. In order to understand better this behavior, we studied a two machine line characterized with single failure mode geometric machines as seen in Figure 4.10. The goal of the analysis is to identify the main factors affecting $V(N)$. The first factor determines the position of throughput bottleneck machine, whereas the second factor determines the position of the machine with higher v. For each possible combination we analytically plot three instances of each case of V as a function of N that varies from 2 to 200. Results of the experiment are reported in Figures 4.11 and 4.12 and Table 4.3.

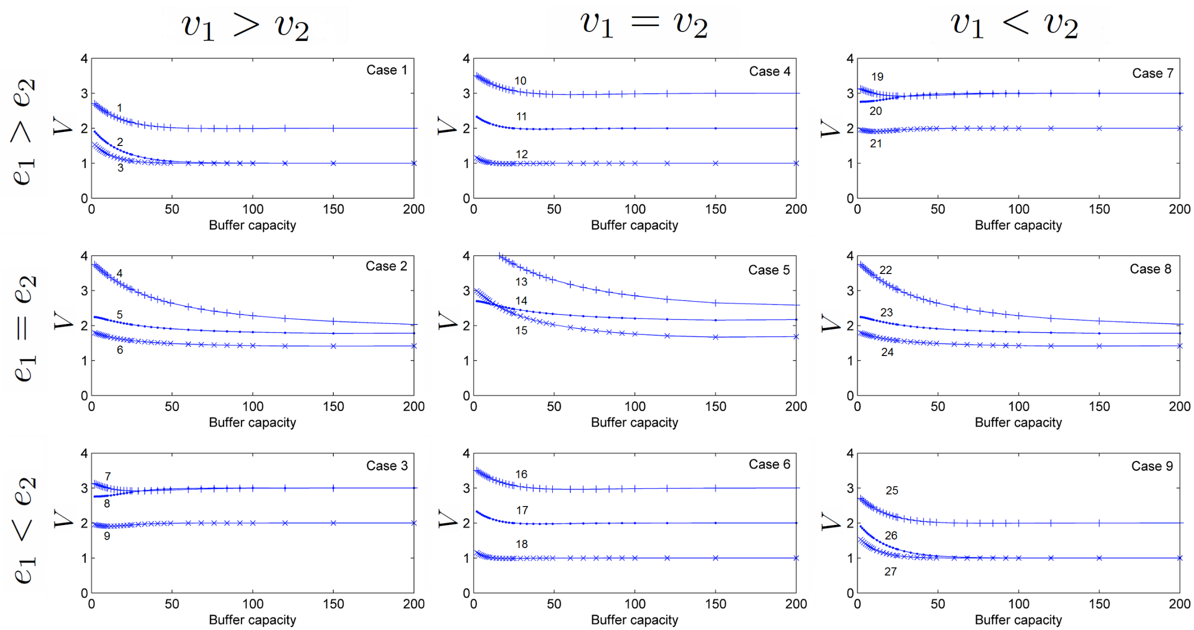


Figure 4.11: The asymptotic variance rate V for the nine possible cases of a two machine line as a function of the buffer capacity. Each output curve represents machine parameters instance that can be found in Table 4.3
.

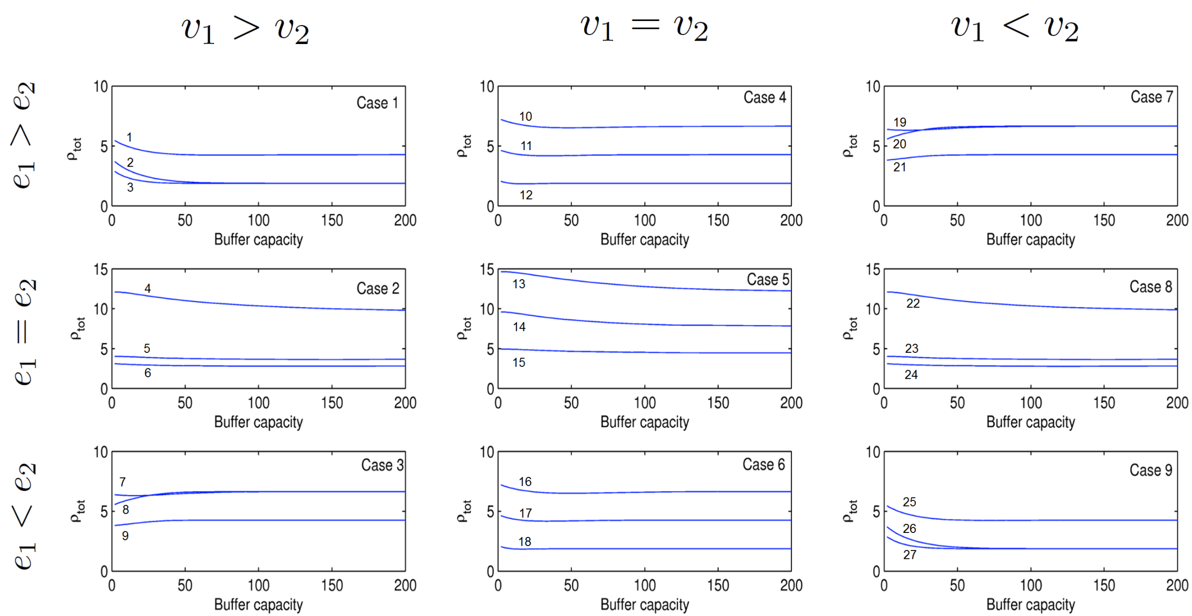


Figure 4.12: The ρ_{tot} for the nine possible cases of a two machine line as a function of the buffer capacity. Each output curve represents machine parameters instance that can be found in Table 4.3
.

Before analyzing single cases of Figure 4.11, looking at equation (4.14) applied to the two machine case:

$$V = E(1 - E)(1 + 2\rho_{tot})$$

and let's define $\tilde{V} = E(1 - E)$ as the variance of the throughput. It is possible to notice that the behavior of $\tilde{V}$ is a concave parabola of the system's throughput E. It is also known that E is a non-decreasing function of the buffer capacity N [29]. We can observe two different effects of N on $\tilde{V}$: if $E(N) < 0.5$ any increase in N will cause E and $\tilde{V}$ to increase, while if $E(N) > 0.5$ any increase in N will cause E to increase, and $\tilde{V}$ to decrease.

The effect of the second term i.e. $(1 + 2\rho_{tot})$ can be seen from Table 4.3 and Figure 4.12. The total autocorrelations ρ_{tot} depends on the eigenvalues describing the system. In some cases ρ_{tot} increases with N like instances (8, 9, 20 and 21) in other cases it decreases with N like (2, 4, 15 and 26). More complex situations show a decreasing then an increasing behavior like (7, 16, 18 and 19).

The results obtained, show that V can be decreasing or increasing function of N, this result is in accordance with results of Tan [79] and Carrascosa [11], whereas it contradicts with the findings of Hendricks [42] who noticed that the V always decreases when the buffer capacity increases.

More insights can be obtained from the nine possible combination of machine parameters presented in Figure 4.11:

- When increasing the size of N, V approaches v of the bottleneck machine, for machines with unequal efficiencies in isolation (unbalanced machines).

- The rate of convergence of V is high for the unbalanced machines and slow for balanced machines.

- We also notice an exact same behavior of V as a function of N between a system and its reverse, like in cases (1,2,3,4) vs cases (9,8,7,6) respectively. This result was also observed by Hendricks [42] for reliable machines and exponentially processing times.

Table 4.3: The values of V, E and ρ_{tot} for the studied 27 experiments at buffer levels N = 5, 20, 45 100 and 200.

Case ID	Instance	e_1	v_1	e_2	v_2	N	5	20	45	100	200
						V	2.5861	2.2286	2.0298	1.9935	1.9998
	1	0.9	3	0.7	2	E	0.6585	0.6821	0.6951	0.6997	0.7000
						ρ_{tot}	5.2500	4.6385	4.2882	4.2435	4.2614
1						V	1.7589	1.3268	1.0750	1.0010	1.0000
$e_1 > e_2$	2	0.9	3	0.7	1	E	0.6591	0.6833	0.6958	0.6998	0.7000
$v_1 > v_2$						ρ_{tot}	3.4143	2.5656	2.0395	1.8824	1.8809
						V	1.4056	1.1052	1.0049	0.9996	0.9936
	3	0.9	2	0.7	1	E	0.6627	0.6894	0.6985	0.7000	0.7000
						ρ_{tot}	2.6441	2.0809	1.8857	1.8798	1.8659
						V	3.6200	3.1410	2.6952	2.2808	2.0313
	4	0.9	3	0.9	2	E	0.8257	0.8486	0.8661	0.8806	0.8891
						ρ_{tot}	12.0775	11.7250	11.1170	10.3436	9.8006
2						V	2.2200	2.0653	1.9250	1.8124	1.7458
$e_1 = e_2$	5	0.7	3	0.7	2	E	0.5669	0.6260	0.6575	0.6780	0.6883
$v_1 > v_2$						ρ_{tot}	4.0209	3.9107	3.7740	3.6509	3.5686
						V	1.7489	1.6014	1.4979	1.4250	1.3812
	6	0.7	3	0.7	1	E	0.5725	0.6348	0.6641	0.6819	0.6905
						ρ_{tot}	3.0729	2.9538	2.8573	2.7849	2.7315
						V	3.0712	2.9298	2.9356	2.9941	3.0000
	7	0.7	3	0.9	2	E	0.6607	0.6865	0.6974	0.6999	0.7000
						ρ_{tot}	6.3497	6.3061	6.4549	6.6277	6.6428
3						V	2.7620	2.8478	2.9705	3.0001	3.0000
$e_1 < e_2$	8	0.7	3	0.9	1	E	0.6661	0.6937	0.6995	0.7000	0.7000
$v_1 > v_2$						ρ_{tot}	5.7090	6.2009	6.5663	6.6432	6.6424
						V	1.9287	1.9294	1.9889	2.0000	2.0000
	9	0.7	2	0.9	1	E	0.6679	0.6952	0.6998	0.7000	0.7000
						ρ_{tot}	3.8479	4.0525	4.2332	4.2619	4.2503
						V	3.4105	3.1222	2.9749	2.9820	2.9995
	10	0.9	3	0.7	3	E	0.6580	0.6809	0.6943	0.6996	0.7000
						ρ_{tot}	7.0773	6.6852	6.5086	6.5941	6.6415
4						V	2.2414	2.0244	1.9759	1.9976	2.0000
$e_1 > e_2$	11	0.9	2	0.7	2	E	0.6616	0.6879	0.6979	0.7000	0.7000
$v_1 = v_2$						ρ_{tot}	4.5053	4.2143	4.1864	4.2558	4.2619
						V	1.0807	0.9873	0.9978	1.0000	1.0000
	12	0.9	1	0.7	1	E	0.6704	0.6967	0.6999	0.7000	0.7000
						ρ_{tot}	1.9455	1.8362	1.8752	1.8807	1.8809
						V	4.3758	3.8805	3.3690	2.8487	2.5475
	13	0.9	3	0.9	3	E	0.8246	0.8452	0.8624	0.8777	0.8872
						ρ_{tot}	14.6243	14.3311	13.6944	12.7728	12.2299
5						V	2.6787	2.5191	2.3515	2.2039	2.1329
$e_1 = e_2$	14	0.7	3	0.7	3	E	0.5628	0.6186	0.6515	0.6743	0.6862
$v_1 = v_2$						ρ_{tot}	4.9434	4.8387	4.6785	4.5177	4.4523
						V	2.8753	2.4384	2.0687	1.7566	1.6186
	15	0.9	2	0.9	2	E	0.8274	0.8530	0.8704	0.8837	0.8910
						ρ_{tot}	9.5664	9.2227	8.6688	8.0448	7.8340

continued on the next page ...

Table 4.3 – Continued

Case ID	Instance	e_1	v_1	e_2	v_2	N	5	20	45	100	200
6 $e_1 < e_2$ $v_1 = v_2$	16	0.7	3	0.9	3	V	3.4105	3.1222	2.9751	2.9823	3.0005
						E	0.6580	0.6809	0.6943	0.6996	0.7000
						ρ_{tot}	7.0773	6.6853	6.5091	6.5948	6.6440
	17	0.7	2	0.9	2	V	2.2414	2.0244	1.9760	1.9978	2.0000
						E	0.6616	0.6879	0.6979	0.7000	0.7000
						ρ_{tot}	4.5053	4.2143	4.1866	4.2562	4.2618
	18	0.7	1	0.9	1	V	1.0807	0.9873	0.9978	1.0001	1.0001
						E	0.6704	0.6967	0.6999	0.7000	0.7000
						ρ_{tot}	1.9455	1.8362	1.8752	1.8811	1.8812
7 $e_1 > e_2$ $v_1 < v_2$	19	0.9	2	0.7	3	V	3.0712	2.9298	2.9354	2.9937	3.0000
						E	0.6607	0.6865	0.6974	0.6999	0.7000
						ρ_{tot}	6.3497	6.3060	6.4545	6.6267	6.6427
	20	0.9	1	0.7	3	V	2.7620	2.8479	2.9703	2.9997	3.0000
						E	0.6661	0.6937	0.6995	0.7000	0.7000
						ρ_{tot}	5.7090	6.2010	6.5659	6.6422	6.6428
	21	0.9	1	0.7	2	V	1.9287	1.9294	1.9889	1.9999	2.0000
						E	0.6679	0.6952	0.6998	0.7000	0.7000
						ρ_{tot}	3.8479	4.0525	4.2331	4.2618	4.2619
8 $e_1 = e_2$ $v_1 < v_2$	22	0.9	2	0.9	3	V	3.6200	3.1410	2.6954	2.2817	2.0389
						E	0.8257	0.8486	0.8661	0.8806	0.8891
						ρ_{tot}	12.0775	11.7252	11.1178	10.3480	9.8362
	23	0.7	2	0.7	3	V	2.2200	2.0653	1.9252	1.8122	1.7457
						E	0.5669	0.6260	0.6575	0.6780	0.6883
						ρ_{tot}	4.0209	3.9107	3.7742	3.6506	3.5685
	24	0.7	1	0.7	3	V	1.7489	1.6014	1.4981	1.4249	1.3812
						E	0.5725	0.6348	0.6641	0.6819	0.6905
						ρ_{tot}	3.0729	2.9539	2.8579	2.7848	2.7315
9 $e_1 < e_2$ $v_1 < v_2$	25	0.7	2	0.9	3	V	2.5861	2.2286	2.0299	1.9941	1.9998
						E	0.6585	0.6821	0.6951	0.6997	0.7000
						ρ_{tot}	5.2500	4.6385	4.2885	4.2449	4.2614
	26	0.7	1	0.9	3	V	1.7590	1.3268	1.0752	1.0014	1.0000
						E	0.6591	0.6833	0.6958	0.6998	0.7000
						ρ_{tot}	3.4144	2.5657	2.0398	1.8832	1.8808
	27	0.7	1	0.9	2	V	1.4056	1.1052	1.0049	0.9999	0.9938
						E	0.6627	0.6894	0.6985	0.7000	0.7000
						ρ_{tot}	2.6441	2.0810	1.8859	1.8807	1.8662

4.2.1 Two Machine Bernoulli Line

The autocorrelation of the Bernoulli machine is known to be equal zero, as explained in Subsection 4.1.5. However, the autocorrelation of Bernoulli machines in a production line is not equal to zero. This section will give more insight on the Bernoulli machine model output variability.

Since the output variability of an isolated Bernoulli machine is zero, we identify three different configurations of the two machine line: the first machine has a higher e, a balanced line, the second machine has a higher e. Figure 4.13 show the asymptotic variance rate V and the throughput E and the total autocorrelations ρ_{tot}. The figure shows that the total autocorrelation coefficient of the transfer lines composed of unbalanced Bernoulli machines approaches *zero* as $\mathrm{N} \to \infty$, making V a function of E only. While for balanced lines we it can be seen from the same figure that $\rho_{tot} \neq 0$ as $\mathrm{N} \to \infty$, meaning that the output of the line is not an independent process anymore. The balanced case studied in Figure 4.13 has a negative autocorrelation structure which turns out to be have a positive effect on V as it makes it decrease.

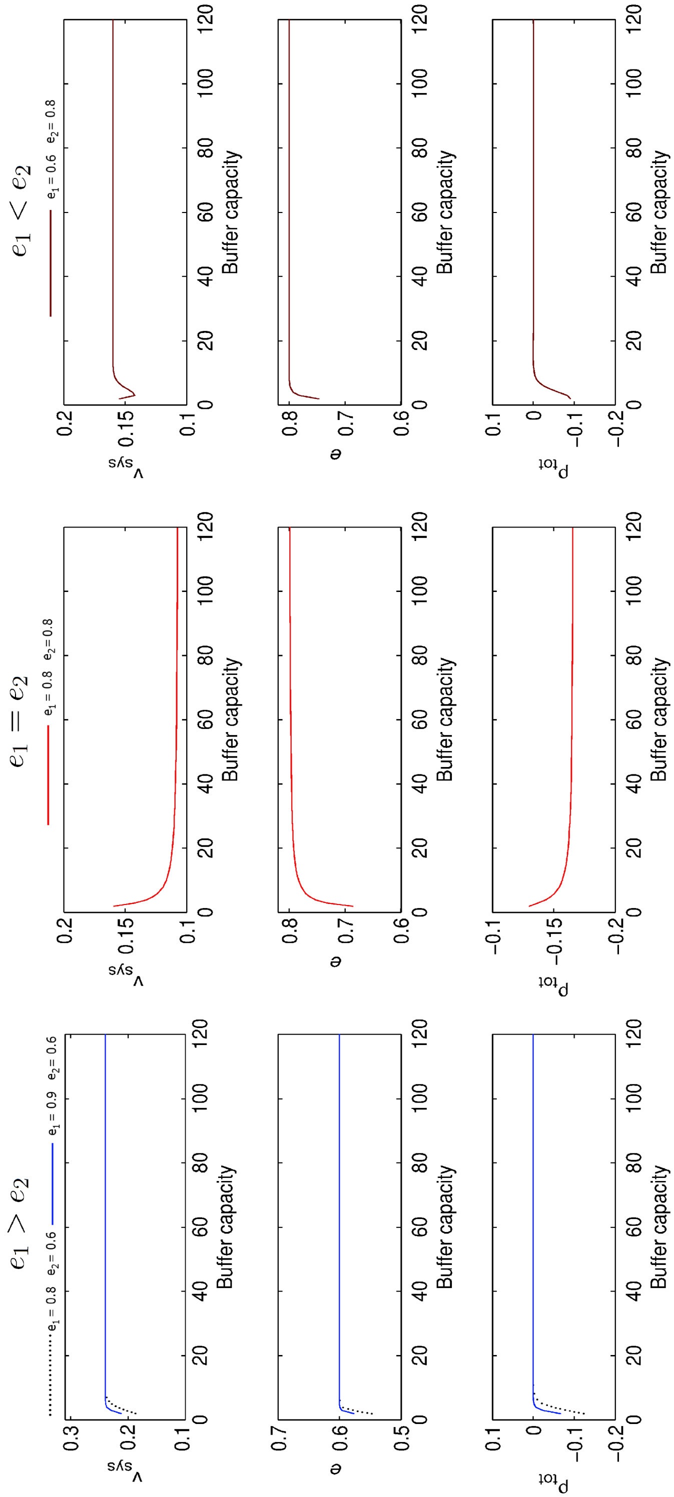


Figure 4.13: The asymptotic variance rate, the throughput and the total autocorrelations as a function of the buffer capacity for a two machine line with Bernoulli machines

4.3 Small Multi-Stage Production Lines

The analytical method proposed in this chapter can still be applied to multi-stage production lines with limited buffer capacities. To apply this approach, the transition probability matrix $\mathbf{P}$ and the reward vector $\boldsymbol{\mu}$ that represent the system should be constructed. Then using equations (4.11) and (4.14) the system's output asymptotic variance rate V can be calculated.

The problem with this approach though, is the state explosion phenomena. The system increase exponentially with the increasing complexity. This exponential growth cause some systems to require infinite times to be solved (given the computer has enough memory). Table 4.4 shows the average evaluation time of a system composed from different single failure mode geometric machines with equal intermediate buffers. Each experiment corresponds to 10 runs for randomly selected machine parameters. It should also mentioned that using a sparse representation of the state space and eliminating transient states can increase the speed of the method significantly. The results were obtained using an Intel Core2 Duo 1.6GHz computer with 3 GB of RAM. It could be seen that the speed of the method depends on the matrix multiplication operation speed. Using the test cases presented in Table 14.4 an estimation of the systems' evaluation time (T_{ev}) in seconds, can be obtained by as a function of non transient states S and the number of significant lag terms k^* as:

$$T_e v = 21.7 + 2.3797{\times}10^{-7} k^* \cdot S^2 - 0.0509 k^* \tag{4.43}$$

with $R^2 = 99.9\%$ indicating that this models completely fits the evaluation time. Finally, Table 4.5 shows the evaluation of different production lines with different lengths and machine parameters (single failure mode Geometric machines) using the proposed and Tan's [84] approaches. Results show that the proposed approach is most cases is slower than Tan's approach. However the last case of Table 4.5 shows that the proposed method was able to solve the matrix analytically where Tan's method could not find a solution with the machine's memory capacity.

4.4 Conclusions

This chapter proposed a methodology to calculate the different output variability indicators for single and small multi-stage manufacturing

Table 4.4: The time needed to evaluate a production line with single failure geometric machine and equal intermediate buffer capacity

Number of machines and buffers	Number of non-transient states	Mean k^*	Mean Evaluation time [s]
case M2-B09	32	600.7	0.206
case M2-B10	36	521.7	0.201
case M3-B02	8	224.0	0.121
case M3-B03	32	408.9	0.260
case M3-B04	72	594.3	0.748
case M3-B05	128	426.7	1.417
case M3-B06	200	413.3	2.783
case M3-B07	288	927.4	13.006
case M3-B08	392	477.5	12.441
case M3-B09	512	694.4	30.640
case M3-B10	648	500.8	35.738
case M3-B11	800	1786.3	185.437
case M3-B12	968	1058.0	163.459
case M4-B02	16	268.6	0.225
case M4-B03	128	332.1	1.394
case M4-B04	429	690.6	28.578
case M4-B05	1024	517.2	124.877
case M4-B06	2000	488.3	458.755
case M4-B07	3456	944.7	2664.638
case M5-B02	32	114.7	0.234

Table 4.5: The output performances for a set of production lines with different lengths, buffer capacities and machine parameters

			Proposed Method				Applying Tan's method [84]		
Line length	Isolated machine's performances (e, v)	buffers capacity	E	V	Number of non-tran -sient states	Evaluation time[s]	E	V	Evaluation time[s]
3	(0.9,10)	3-5	0.7530	17.212	64	2.01	0.7533	17.182	0.33
3	(0.95,2)	8-8	0.8753	4.168	392	66.43	0.877	4.114	0.37
4	(0.7,3)	2-2-2	0.3756	1.717	16	0.05	0.3841	1.686	0.21
4	(0.9,8)	2-2-2	0.6927	14.527	15	0.08	0.6931	14.476	0.25
4	(0.8,4)	4-3-2	0.5110	3.879	96	1.25	0.5198	3.812	0.56
4	(0.9,8)	4-4-4	0.6974	14.319	390	104.98	0.6994	14.324	2.14
4	(0.7,3)	4-4-4	0.4073	1.789	432	25.09	0.4230	1.770	1.85
4	(0.8,4)	10-5-3	0.5324	3.904	1152	442.10	0.5495	3.815	9.62
5	(0.8,4)	2-2-2-2	0.4484	3.384	32	0.09	0.4556	3.308	3.92
5	(0.8,4)	4-4-4-4	0.4714	3.433	2480	3858.20	Out of memory		

systems with general Markovian structure and binary reward. The analysis shows the importance of considering the autocorrelation structure while estimating the output variability of the system. The approach is general and it can be applied to several different machine models. Results show relevant, previously uninvestigated, relations between the output variance and the machine reliability parameters and the buffer sizes. In particular, depending on the machine parameters, reducing the MTTR or increasing the MTTF of the machine may even have a negative impact on the output variability. This counter intuitive result is important for choosing improvement options that will have positive effect on both e and v. Moreover, it was shown that increasing the buffer size may reduce or increase the output variability, and an explanation for this behavior was drawn.

Chapter 5

Evaluation of the Output Variability in Long Multi-Stage Production Lines

5.1 Introduction

The exact method proposed in Section 4.3 could be in principle applied to deliberately long manufacturing systems. However, as the dimension of the state-space, S, consistently increases, it is practically inapplicable for more than 4 machines. Therefore, this section introduces four approximate methods based on the decomposition and aggregation approaches that enable to evaluate the variance rate for any deliberately long production line.

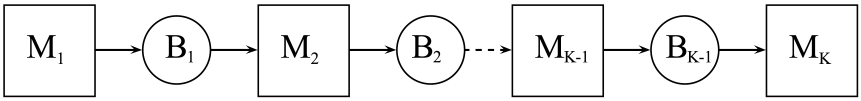


Figure 5.1: A production line with K machines

5.2 System Description and Assumptions

The considered discrete manufacturing system is composed of K unreliable machines separated by $K-1$ dedicated buffers with limited capacity. The i^{th} machine and its dedicated buffer are denoted with M_i and B_i, respectively. B_i has capacity equal to N_i and it contains pieces already processed by M_i. Machine M_i is said to be blocked if the downstream B_i is full. Similarly, M_i is said to be starved if the upstream buffer B_{i-1} is empty. The first machine is never starved and the last machine is never blocked. The system is schematically represented in figure 5.1, where squares and circles represent machines and buffers respectively. Further assumptions are listed below:

- Time is discrete and all the machines have equal and constant processing times. Time is scaled so that the processing cycle of each machine takes exactly one time unit.
- Machines start their operations at the same time period, thus the system is synchronous.
- Whenever a machine M(i) begins to process a workpiece, there is a constant probability $p_{j,i}$ that fails in mode j. Time between failures (TBF) follows a geometric distribution with mean $1/p_{j,i}$.
- Whenever a machine is failed in a given mode j, there is a constant probability $r_{j,i}$ that will be repaired from failure of type j. Time to repair (TTR) follows a geometric distribution with mean $1/r_{j,i}$.
- Failures are operation dependent, when a machine is not processing a workpiece (i.e., it is starved or blocked) can not fail.
- By convention, repairs and failures occur at the beginning of time units and changes in buffer levels take place at the end of the time units.
- Blocking before service is considered.

The assumptions presented in Chapter 3 are applied here as well. The only difference is that the proposed methods (that are going to be explained later) were developed considering multiple failure modes geometric machines. However, if the studied machines are modeled with a General Markovian structure, equations (4.33) can be used to transform the machine into an equivalent single failure mode geometric machine that has the same e and v, thus the proposed methods can be still applied.

The objective of this chapter is to derive methods to approximately estimate the production rate E and the asymptotic variance rate V for multi-stage production system described by the previous assumptions.

5.3 The CMT method

5.3.1 Description of the Method

The CMT method named after Colledani, Matta, and Tolio [19], this method uses the idea of decomposition originally proposed in Gershwin [29] for evaluating the output variability for long multi-stage production systems. The production line can be approximately represented by a set of small production lines, also named building blocks. Building blocks are easier to analyze because of their lower complexity compared to that of the original system. Specifically, the approach is to decompose the K-machine system into a set of $K-1$ two-machine one-buffer sub-systems, i.e. one for each buffer in the original system. Each building block is denoted as $BB(k)$ and is characterized by one upstream pseudo-machine $M^u(k)$, one downstream pseudo-machine $M^d(k)$ and one buffer $B(k)$. Thus a line with four machines such that in Figure 5.2 is decomposed into three building blocks.

The coherence among building blocks is made possible by the definition of decomposition equations that establish proper relationships among them. Parameters of building blocks have to be defined so that each one represents the behavior of the original system and in order to respect the equivalence between the flows passing through the buffers of the original line, and through the buffers in the building blocks. The parameters to set are the values of the variables assigned to each building block that rule the interruption of the flow of parts (failure and repair probabilities) in each buffer. Specifically, all the

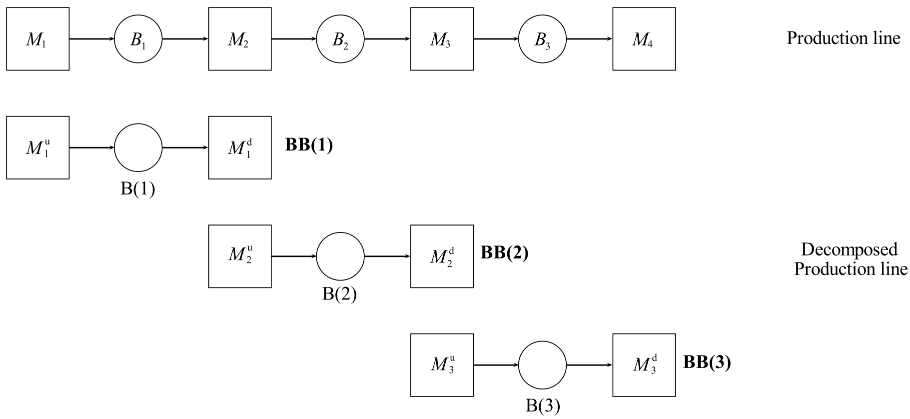


Figure 5.2: The decomposition of a four machine line

interruptions of the material flow entering (leaving) the buffer B_k are modeled by the pseudo-machine $M_u(k)$ ($M_d(k)$). Similarly to the decomposition method proposed in [85] for lines with multiple failure mode machines, we assign to every pseudo-machine remote states modeling the propagation of starvation and blocking in the system.

The average production rate for each building block i is denoted with $E(i)$, $Ps_j^u(i)$ is the probability of starvation of machine $M^d(i)$ due to the remote failure mode j and $Pb_j^d(i)$ is the probability of blocking of machine $M^u(i)$ due to the remote failure mode j. In order to find the new remote failure probability of a down stream machine $p_j^u(i)$ the flow-rate idle time equation is applied:

$$p_j^u(i) = \frac{Ps_j^u(i-1)}{E(i-1)} \cdot r_j^u(i) \tag{5.1}$$

This condition guarantees the propagation of the first moment of the output [29]. The opposite is applied to find the remote failure probabilities due to blocking for downstream machine of $BB(i)$ as:

$$p_j^d(i-1) = \frac{Pb_j^d(i)}{E(i-1)} \cdot r_j^d(i) \tag{5.2}$$

Assuming to have solved the production line by means of the decomposition technique, the last building block can be analyzed as described in section 4.1.2, for example the last building block in the decomposed line in Figure 5.2 is BB(3). The first machine of the last building block represents the portion of the line that is upstream the last buffer. The behavior of this machine is made coherent with that of the corresponding real system by means of the additional remote failures, that model the interruptions of the material flow entering $B(K-1)$. The parameters of these additional failures are the solutions of decomposition equations. The second machine of the last building block is exactly as the last machine of the production line. Therefore, from the decomposition technique the parameters of the last building block are obtained and equations (4.11) and (4.14) can be applied. Note that the calculated asymptotic variance rate is an approximation of the true value, since the behavior of the last building block does not model exactly that of the real production line.

5.3.2 CMT Algorithm

The DDX Algorithm [see [22]] has been developed to solve decomposition equations. This algorithm consists of the following main steps:

- STEP 0: Initialization. The pseudo-machine local failure parameters are initialized to the failure and repair parameters of the original line. Pseudo-machines remote failure parameters are initialized to some given values $\lambda = 0.05$.

- STEP 1: Upstream pseudo-machine failures updating: for $i = 2, \ldots, K-1$ decomposition equation (5.2) is solved and the $BB(i)$ is evaluated.

- STEP 2: Downstream pseudo-machine failure updating: for $i = K-2, \ldots, 1$ decomposition equation (5.2) is solved. $BB(i)$ is evaluated.

- STEP 3: repeat STEPS 1 and 2 until no further changes in the reliability parameters of the machines are verified. In this situation, the conservation of flow equation must be met:

  $$E(i-1) = E(i) \tag{5.3}$$

  for $i = 2, \ldots, K-1$.

- STEP 4: Evaluate the output measures. The CMT method calculates the output variability measure V from the last building block of the decomposed line using the proposed approach presented in Chapter 4

5.3.3 Modified CMT method

In the CMT method, the line is decomposed into two machine building blocks that assure that the material flow through each building block is equal to that of the line. The idea of this modified method (let's calls it CMT3) is to decompose the line into normal two machine building blocks only keeping the last building block composed of "a three machine two buffer system". By this approach the last part of the line is more accurately represented. However, this approach will cause the state explosion phenomena if the buffer capacities among the last three machine building block are high. The method can be used efficiently for long production lines with low buffer capacities, especially when the last two buffers have low capacities.

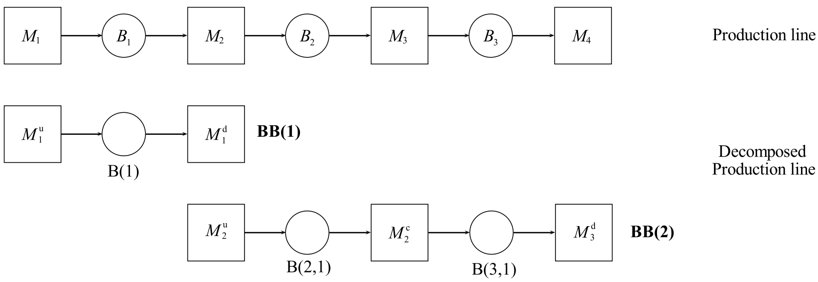


Figure 5.3: The decomposition of a four machine line

The evaluation of such method is not different from the normal CMT, the only difference lies in the way the lines is decomposed. The algorithm of CMT method can be applied for this method as well.

5.4 Machine Aggregation Method

Aggregation methods try to represent a complex system (in terms of number of states) with a simpler one. Such methods have been proposed to production lines by Lim et al. [54], however the current method differs in the aggregation equations used.

The idea of Aggregation method (AGG) is to start from the beginning of the line, considering only a sub-system of two machines (M_1 and M_2) and the intermediate buffer between them. The output coming out of M_2 represents the output of the subsystem, thus one can replace the subsystem with a machine M_{eq} that will have the same output measures of the subsystem i.e. average throughput and output asymptotic variance rate.

If the two machines are single failure mode machines, and the buffer between them has the capacity of 10, then the total number of states is $2 \times 2 \times 11 = 44$ states. The idea of the method is to find a geometric machine with only two states that has the same output performances of the original two machine system. To find the parameters of the equivalent machine M_2 which are (p_{eq} and r_{eq}) the following equations should be applied:

$$p_{eq} = \frac{2\,e\,(e-1)^2}{-e^2+e+v} \tag{5.4}$$

$$r_{eq} = \frac{2 \cdot e^2 \cdot (e-1)}{(e^2-e-v)} \tag{5.5}$$

which were derived from the closed formula of the asymptotic variance rate and the average throughput of the single failure mode geometric machine [Gershwin [28]].

5.4.1 Solution Algorithm

To apply the aggregation method, the following steps should be carried out:

- STEP 1: Evaluate the first subsystem i.e. the first two-machine one-buffer system. The evaluation can be done by any analytical method that calculates the first two moments of the output.
- STEP 2: Transform the subsystem into a geometric machine with parameters p_{eq} and r_{eq}. After evaluating the first subsystem, a new equivalent machine M_{eq} can be found having the same output performances of the original line using equations (5.4) and (5.5).
- STEP 3: Evaluate the new subsystem consisting of the first M_{eq} and the following buffer i.e buffer 2 and machine 3.
- STEP 4: Repeat steps 2 and 3 until the end of the line, as depicted in Figure 5.4.

5.4.2 Limitations of the Aggregation Approach

This approach is applicable for a wide range of values of e and v. However, the critical region where it will not work is when v is less

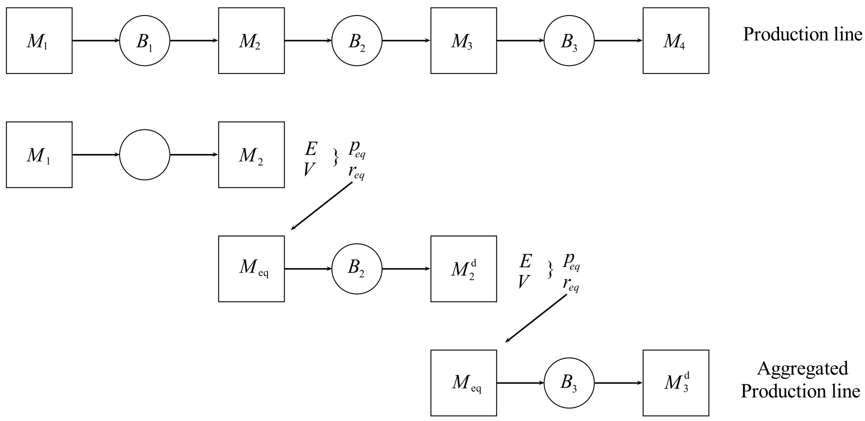


Figure 5.4: The application of the Aggregation method for the estimation of V for a multi-stage manufacturing line

than 0.1, Figure 5.5 shows the values e and v where this approach does not work.

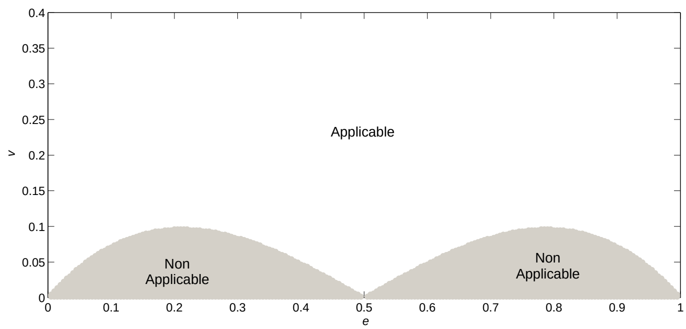


Figure 5.5: The applicability of the aggregation method as a function of e and v

5.5 AC-Decomposition Method

Differently from the CMT method, the derivation of the failure and repair probabilities for the remote failures in order to convey both the first and the second moments of the asymptotic production between building blocks. This innovative feature allows higher accuracy in the output variance estimation.

The basic idea is to use the exact method presented in the Chapter 4.3 to study the building block output mean and variance. Then, the upstream pseudo-machine parameter of the next two-machine line are updated in order to match both the mean and the variance of the output of the previous building block. By applying a forward and backward pattern, the method reaches convergence and both the asymptotic mean and variance of the output can be estimated. The equations for propagating the moments are derived in the next subsections.

5.5.1 Analysis of the Subsystem

The proposed method benefits from the results presented in Ciprut et al. [14]. In their work, a closed-form formula for calculating the asymptotic output mean and variance for an unreliable single machine with generally distributed up and down times was derived by using a fluid Markovian model. More recently, the discrete time version of that formula was derived in [44]. Defining m_D and m_U as the mean down and up times of the machine, cv_D^2 and cv_U^2 as the square coefficient of variation of down and up times of the machine, and I as the unavailability factor, i.e. ${}^{m_D}\!/_{m_U}$, the following can be written:

$$e = \frac{1}{1+I} \tag{5.6}$$

$$v = \left(cv_U^2 + cv_D^2\right)\left(\frac{m_D I}{(1+I)^3}\right) \tag{5.7}$$

For multiple failure mode machines of the type of those considered in this paper, featuring geometrically distributed failure and repair times, we can write:

$$m_U = \frac{1}{P} \tag{5.8}$$

$$m_D = \sum_{\forall j} \frac{p_j}{P}\left(\frac{1}{r_j}\right) \tag{5.9}$$

where P is the sum of all failure probabilities and it is equal to $\sum_{\forall j} p_j$. The cv^2 for the up and down times can be calculated as:

$$cv_U^2 = 1 - P \tag{5.10}$$

$$cv_D^2 = \frac{\sum_{\forall j} \frac{p_j}{P} \frac{2-r_j}{r_j^2} - m_D^2}{m_D^2} \tag{5.11}$$

Substituting equations (5.8, 5.9, 5.10 and 5.11) in equations (5.7) and (5.6) we obtain e, v as:

$$e = \frac{1}{1 + \sum_{\forall j} I_j} \tag{5.12}$$

$$v = \frac{\sum_{\forall j} I_j \left(\frac{2-r_j}{r_j}\right) - \left(\sum_{\forall j} I_j\right)^2}{(1 + \sum_{\forall j} I_j)^3} \tag{5.13}$$

where I_j is the unavailability factor generated by the j failure mode and it is equal to p_j/r_j. By substituting equation (5.12) into (5.13) the following expression can be obtained:

$$L = \frac{v}{e^3} + \frac{1-e}{e} = \sum_{\forall j} I_j \left(\frac{2-r_j}{r_j}\right) \tag{5.14}$$

These equations will be used to derive the decomposition equations.

5.5.2 Pseudo-Machine States

Here the detailed analysis is reported for the parameters of the pseudo-machines of the building blocks $BB(i)$, with $i = 2, .., K-1$. Similar equations will be given for the machines of the building block $BB(i-1)$ in the algorithm derivation section.

In order to mimic the disruption regulating the dynamics of the material flow entering buffer B_i in the original line, the pseudo-machine $M^u(i)$ must be a multiple failure mode machine featuring two sets of failures: (i) local failures, that model the stops in the material flow due to failure of the machine M_i that is just upstream the buffer B_i and,(ii) one remote failure that models the stops in the material flow entering B_i due to a failure of upstream machines M_j, $j = 1, .., i-1$ that generate propagation of starvation. The failure and repair probabilities of local failure for machine $M^u(i)$ are simply obtained by copying the failure and repair probabilities of the corresponding machine in the original line M_i. However, the failure and repair probabilities of the remote failure, $p^u(i)$ and $r^u(i)$, are unknown and need to be calculated with properly defined decomposition equations, that allow to match the first two moments of the output production of the previous building block $BB(i-1)$. Figure 5.6 depicts a decomposed three machine line. It shows green circles as up states, single boundary red circles as local failure states and double boundary red circles as remote failure states. Figure 5.7 depicts an example of a Pseudo up-stream machine of BB(i), where subscript r refers to the remote failure or repair mode and subscript l refers to a local failure/repair rate.

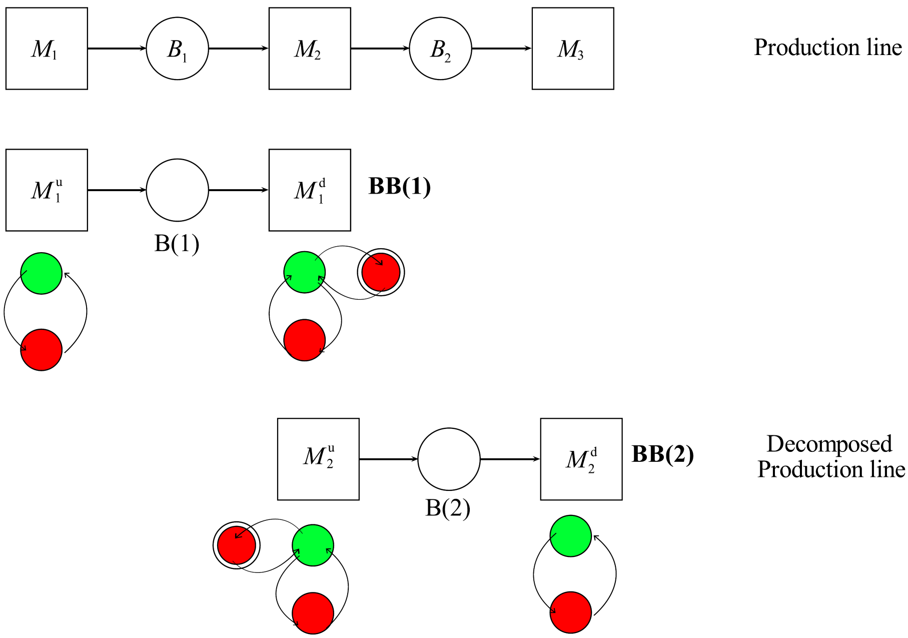


Figure 5.6: The decomposition of a three machine line using the AC decomposition method

5.5.3 Conservation of the Moments of the Material Flow

In [29] the conservation of flow equation was introduced. This equation stated that in a serial production lines the asymptotic throughput is conserved. In other words, observing the material flow from any machine in the system, the same average throughput is detected, which is also the average throughput of the line. This property of serial lines is a fundamental property for the traditional decomposition methods that is also strictly connected to the reversibility property of production lines [63]. The discussion in Section 4.2 indicates that the same arguments hold also for what concerns the asymptotic variance of the output. Therefore, in the long run, the output variability observed at different

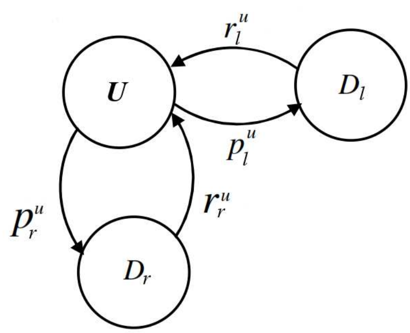


Figure 5.7: The pseudo machine states for a upstream machine in a BB(i)

observation points in the system is the same. We make use of this property for the newly developed decomposition equations. Specifically, we assign $p^u(i)$ and $r^u(i)$ to machine $M^u(i)$ such that building block $BB(i)$ has the same asymptotic throughput and variance of the previous building block $BB(i-1)$, computed with the method proposed in this paper. Similarly, we assign $p^d(i)$ and $r^d(i)$ to pseudo-machine $M^d(i)$ such that building block $BB(i)$ has the same asymptotic throughput and variance of the next building block $BB(i+1)$. By iteratively updating the parameters of the upstream pseudo-machines in the forward phase of the algorithm and of the downstream pseudo-machines in the backward phase of the algorithm, the property of conservation of the moments of the material flow is met.

5.5.4 Decomposition Equations

The unknown parameters for the pseudo-machine $M^u(i)$ are two, $p^u(i)$ and $r^u(i)$, thus two conditions are needed to express, one on the average throughput propagation and one on the asymptotic variance propagation. Due to the flow-rate idle time equation [29], the first equation is:

$$p^u(i) = I^u(i)r^u(i) = \frac{Ps(i-1)}{E(i-1)}r^u(i) \tag{5.15}$$

With this condition, the first moment of the output production is propagated. In order to propagate also the second moment, equation (5.14) has to be manipulated to find the value of $r^u(i)$ that makes $M^u(i)$ matching the variance rate V of the building block $BB(i-1)$. Such manipulation leads to the following equation:

$$r^u(i) = \frac{2I^u(i)}{Z^u(i) + I^u(i)} \tag{5.16}$$

where

$$Z^u(i) = L^u(i) - \sum_{j=1}^{F_i} I_{i,j} \frac{2 - r_{i,j}}{r_{i,j}} - I^d(i-1) \frac{2 - r^d(i-1)}{r^d(i-1)} \tag{5.17}$$

and

$$L^u(i) = \frac{V(i-1)}{E(i-1)^3} + \frac{1 - E(i-1)}{E(i-1)} \tag{5.18}$$

Similar equations can also be derived to find the parameters $p^d(i)$ and $r^d(i)$ of the downstream pseudo-machine $M^d(i)$:

$$r^d(i) = \frac{2I^d(i)}{Z^d(i) + I^d(i)} \tag{5.19}$$

where

$$Z^d(i) = L^d(i) - \sum_{j=1}^{F_{i+1}} I_{i+1,j} \frac{2 - r_{i+1,j}}{r_{i+1,j}} - I^u(i+1) \frac{2 - r^u(i+1)}{r^u(i+1)} \tag{5.20}$$

and

$$L^d(i) = \frac{V(i+1)}{E(i+1)^3} + \frac{1 - E(i+1)}{E(i+1)} \tag{5.21}$$

5.5.5 Solution Algorithm

The following iterative procedure allows to properly solve the proposed decomposition equations:

1: STEP0:INITIALIZATION. Decompose the $K-$machine line into $K-1$ sub-systems as shown in Figure 5.2. Initialize the parameters of the local failures of the pseudo-machines to those of the corresponding machines in the original line. Do not consider remote failures at the first iteration of the algorithm. For $i = 1, .., K-1$ evaluate the first and second order performance of the building block $BB(i)$ with the method proposed in Section 4.2.

2: STEP1:FORWARD PHASE. For $i = 2, .., K-1$. Use equations (5.15) and (5.16) to find the unknown remote failure probabilities of the pseudo-machine $M^u(i)$ of $BB(i)$. Evaluate the first and second order performance of the building block $BB(i)$ with the method proposed in Section 4.2.

3: STEP2:BACKWARD PHASE. For $i = K - 2, .., 1$. Use the following equations to find the unknown remote failure probabilities of the pseudo-machine $M^d(i)$ of $BB(i)$.

$$p^d(i) = I^d(i)r^d(i) = \frac{Pb(i+1)}{E(i+1)} r^d(i) \tag{5.22}$$

$$r^d(i) = \frac{2I^d(i)}{Z^d(i) + I^d(i)} \tag{5.23}$$

$$Z^d(i) = L^d(i) - \sum_{j=1}^{F_{i+1}} I_{i+1,j} \frac{2 - r_{i+1,j}}{r_{i+1,j}} - I^u(i+1)\frac{2 - r^u(i+1)}{r^u(i+1)} \tag{5.24}$$

$$L^d(i) = \frac{V(i+1)}{E(i+1)^3} + \frac{1 - E(i+1)}{E(i+1)} \tag{5.25}$$

Evaluate the first and second order performance of the building block $BB(i)$ with the method proposed in section 4.2.

4: STEP3:Repeat steps 1-2 until: $\underset{\forall i \neq j}{max} \triangle |v_{BB(i)} - v_{BB(j)}| < \epsilon$

Finally, since the method approximates the output moments of the BB with the moments of a geometric machine, situations were the BB output moments are far from being geometric could cause algorithmic errors. In particular, resulting remote failure and repair probabilities could be outside their limits $[0, 1]$, at some iteration. To fix this issue if at any step of the algorithm $r^u(i)$ or $r^d(i)$ is outside the $[0, 1]$ limit, its value is recalculated by adjusting the value of Z in the following way:

$$Z^u(i) = L^u(i) - \sum_{j=1}^{F_i} I_{i,j} \frac{2 - r_{i,j}}{r_{i,j}} \tag{5.26}$$

$$Z^d(i) = L^d(i) - \sum_{j=1}^{F_{i+1}} I_{i+1,j} \frac{2 - r_{i+1,j}}{r_{i+1,j}} \tag{5.27}$$

In the following, we will refer to the standard method as AC and to the method with this modification as $\hat{\text{AC}}$. With this approximation the moments are not fully conveyed, but this guarantees obtaining a solution for any combination of first and second moments of the output.

Chapter 6

Numerical Results and System Behavior

6.1 Methodology Validation

In order to test the accuracy and the speed of the proposed methods for calculating the asymptotic variance rate V for long manufacturing systems, the developed approximate analytical methods have been implemented in MATLAB^{TM} version 7.9 (R2009b). The PC used to run the proposed methods is an Intel Core2 Duo 1.6 GHz with a RAM of 3 GB. The developed methods have been widely validated by comparing results with those obtained by a Discrete Event Simulator (DES) also developed on MATLAB. The same PC has been used to run the simulation model too.

The simulation model was developed considering the same assumptions in Chapter 5. This chapter will present the method in which the asymptotic behavior time has been determined (run time), and it will also explain in details the factorial design used for comparison.

In order to obtain statistical significance of the simulation results, 300 repetitions of each case has been performed. Moreover, the transition time window (as will be explained in the following) has been neglected in the calculation of the performance measures. In the following section, the results of different proposed methods will be tested, which is the basis of comparison between the different methods. In the accuracy validation phase, the errors in the estimation of the system performances have been calculated as:

$$\varepsilon_{\text{method}} = \frac{\bar{\phi}_{\text{simulation}} - \phi_{\text{method}}}{\bar{\phi}_{\text{simulation}}} \tag{6.1}$$

where $\bar{\phi}_{\text{simulation}}$ is the average value of the output (E or V) of 300 replicates of each simulation test.

6.1.1 Determining Simulation Warm-Up Period

To determine the run length that would grantee arriving at the steady-state performance the Welch method [90] has been used. The Welch method is a graphical method indicating when the response arrives the steady state. Figures 6.1 and 6.2 show the average V and E for 300 replications and over a time period from 5×10^4 to 1.5×10^6 cycle times. The simulated line is composed of ten machines and nine buffers, furthermore, two alternating machine types M_1 and M_2 have been used, each of these machines has $[e, v]$ equal to [0.9,2] and [0.95,3], respectively, while the buffer used between each pair of machines is equal to 10. Figure 6.2 shows that E reaches the steady state around 5×10^5 cycle times, while it can be seen from Figure 6.1 that the warm-up time for V to arrive to the steady state is around $1.2{\times}10^6$ cycle times. Thus the warm-up period of 1.2×10^6 will be used to study V of the 10 machine line at steady state.

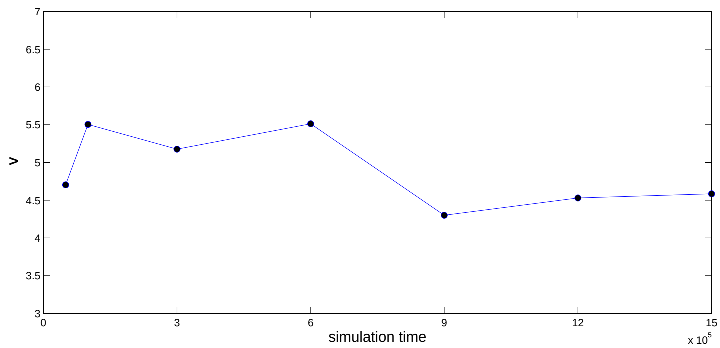


Figure 6.1: The average asymptotic variance rate V of 300 replications of a long manufacturing line at different time periods, for a system of 10 machines-9 buffers line

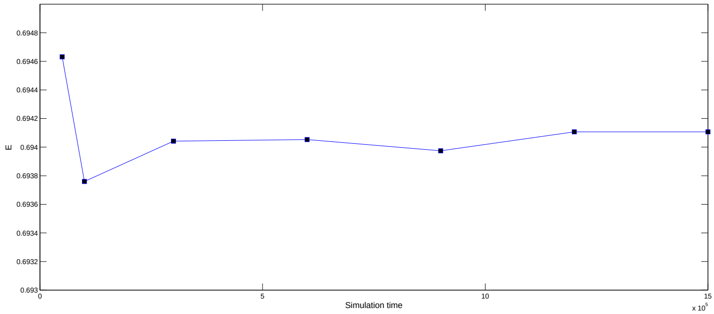


Figure 6.2: The average throughput E of 300 replications of a long manufacturing line at different time periods, for a system of 10 machines for a system of 10 machines-9 buffers line

6.1.2 The Factorial Design

A full non-replicated factorial design was developed and analyzed, considering five factors each of two levels. Table 6.1 shows the high and low levels of the different factors in our experiment. In total, 32 different configurations of production lines were evaluated. The first factor is the line length which sets the number of machines and buffers in the line. The second factor indicates the capacity of the first buffer (B_i) in the line. For the third factor, we use buffer increment factor α to indicate if the line is balanced i.e. equal buffer capacities ($\alpha = 0$) or not. Specifically, buffer $B_{i+1} = B_i + \alpha$. For each simulated configuration, 300 simulation replicates were carried out, each one of 10^6 time units of warm-up and 1.2×10^6 time units, where statistics were collected.

Finally, the last two factors take into account the performance of the machines in isolation, in terms of e and v of the machine's output. The considered responses were the average percentage error and the average percentage absolute error of the method in calculating E and V of the system's output at time $T = 1.2 \times 10^6$ and the computer time required to generate the solution.

Table 6.1: Levels of the factors considered in the experiment

Factor	Low Value	High Value
Line Length / Num. of Buffers	4/3	10/9
Buffer Capacity	3	15
Buffer increment Factor (α)	0	2
Single Machine Efficiency in isolation (e)	85%	95%
Single Machine Asymptotic variance rate (v)	0.361	1.80

The output of each run of the simulation is reported in Table 6.2. Furthermore, it can be seen that the confidence interval for V even after taking 300 replications is wide, see cases: 8,12 and 20.

6.2 Results

6.2.1 Accuracy Testing

Results regarding E and V calculated from the studied methods on the factorial design used are summarized in Table 6.3 and fully shown in Tables 6.7 and 6.8. The accuracy of calculating the mean production rate E is very high in all the methods except for Aggregation. Figure 6.3 shows a Boxplot of the percentage errors against simulation. It can be seen that the order in terms of accuracy (in descending order) is CMT3, CMT, AC, $\hat{AC}$ and finally AGG.

While estimating V, all the methods outperforms the CMT method, as can be seen in Figure 6.4. It can be also noticed that AC and CMT3 method are very accurate in both calculating E and V. However, they both suffer from some limitations; the AC method worked only in 18 out of the 32 experiments performed due to the constraints in the parameters' variability range. On the other hand the CMT3 method was feasible to apply to half of the factorial design (only 4 machine lines) due to the state space explosion phenomena.

On the other hand, both the Aggregation and $\hat{AC}$ methods show to be a trade off solution; the $\hat{AC}$ always provide a result at the cost

Table 6.2: The output of of the DES at $t = 1.2{\times}10^6$ and 300 replications

Case Number	Number of machines	Buffer capacity	Increment factor	e	v	E LCI	E Mean	E UCI	V LCI	V Mean	V UCI	T_{ev}
1	4	3	0	85%	0.361	0.6825	0.6826	0.6827	0.3794	0.4426	0.5230	4866
2	4	3	0	85%	1.8	0.6210	0.6212	0.6214	1.4903	1.7385	2.0547	4772
3	4	3	0	95%	0.361	0.8447	0.8449	0.8450	0.8085	0.9431	1.1146	4960
4	4	3	0	95%	1.8	0.8313	0.8317	0.8320	4.2811	4.9941	5.9023	4937
5	4	3	2	85%	0.361	0.7307	0.7308	0.7309	0.3328	0.3883	0.4589	4942
6	4	3	2	85%	1.8	0.6551	0.6553	0.6555	1.4081	1.6427	1.9414	4816
7	4	3	2	95%	0.361	0.8651	0.8652	0.8654	0.8157	0.9515	1.1246	4984
8	4	3	2	95%	1.8	0.8391	0.8395	0.8398	3.7205	4.3401	5.1294	4944
9	4	15	0	85%	0.361	0.8112	0.8113	0.8113	0.2209	0.2577	0.3046	5023
10	4	15	0	85%	1.8	0.7425	0.7427	0.7429	1.2111	1.4128	1.6697	4929
11	4	15	0	95%	0.361	0.9097	0.9099	0.9100	0.5041	0.5881	0.6950	5030
12	4	15	0	95%	1.8	0.8680	0.8683	0.8687	3.8926	4.5408	5.3666	4974
13	4	15	2	85%	0.361	0.8153	0.8154	0.8154	0.2073	0.2418	0.2858	5026
14	4	15	2	85%	1.8	0.7506	0.7508	0.7510	1.1932	1.3919	1.6450	4956
15	4	15	2	95%	0.361	0.9133	0.9134	0.9135	0.4427	0.5164	0.6104	5023
16	4	15	2	95%	1.8	0.8720	0.8723	0.8726	2.8235	3.2938	3.8928	4993
17	10	3	0	85%	0.361	0.6028	0.6029	0.6030	0.2118	0.2470	0.2920	12551
18	10	3	0	85%	1.8	0.4775	0.4776	0.4778	0.9170	1.0697	1.2642	12042
19	10	3	0	95%	0.361	0.7364	0.7366	0.7368	0.9564	1.1156	1.3185	12690
20	10	3	0	95%	1.8	0.6826	0.6830	0.6833	4.8199	5.6226	6.6451	12455
21	10	3	2	85%	0.361	0.7274	0.7275	0.7276	0.2923	0.3409	0.4030	13003
22	10	3	2	85%	1.8	0.6216	0.6218	0.6220	1.0738	1.2526	1.4804	12593
23	10	3	2	95%	0.361	0.8467	0.8468	0.8470	0.7955	0.9280	1.0968	13108
24	10	3	2	95%	1.8	0.7659	0.7663	0.7666	3.7365	4.3587	5.1514	12773
25	10	15	0	85%	0.361	0.7974	0.7974	0.7975	0.1442	0.1682	0.1988	13252
26	10	15	0	85%	1.8	0.6928	0.6930	0.6931	0.8917	1.0402	1.2294	12876
27	10	15	0	95%	0.361	0.8870	0.8871	0.8872	0.4682	0.5461	0.6454	13269
28	10	15	0	95%	1.8	0.7978	0.7981	0.7984	3.1736	3.7022	4.3754	12950
29	10	15	2	85%	0.361	0.8115	0.8116	0.8117	0.1590	0.1855	0.2192	13358
30	10	15	2	85%	1.8	0.7297	0.7298	0.7300	0.8041	0.9380	1.1086	13044
31	10	15	2	95%	0.361	0.9046	0.9047	0.9048	0.3682	0.4295	0.5076	13383
32	10	15	2	95%	1.8	0.8289	0.8292	0.8295	2.8027	3.2694	3.8640	13078

Table 6.3: Aggregated accuracy results of the different approaches.

		AC	$\hat{AC}$	CMT	CMT3	AGG
E	% Avg. err.	-0.59%	-0.96%	-0.43%	0.42%	3.10%
	% Avg. |err.|	0.911%	1.066%	0.564%	0.42%	3.10%
V	% Avg. err.	-5.03%	-11.08%	-34.39%	-5.25%	-7.94%
	% Avg |err.|	8.781%	16.958%	35.821%	8.90%	11.71%
	% within 95% CI	88.9%	65.63%	34.38%	93.75%	78.13%

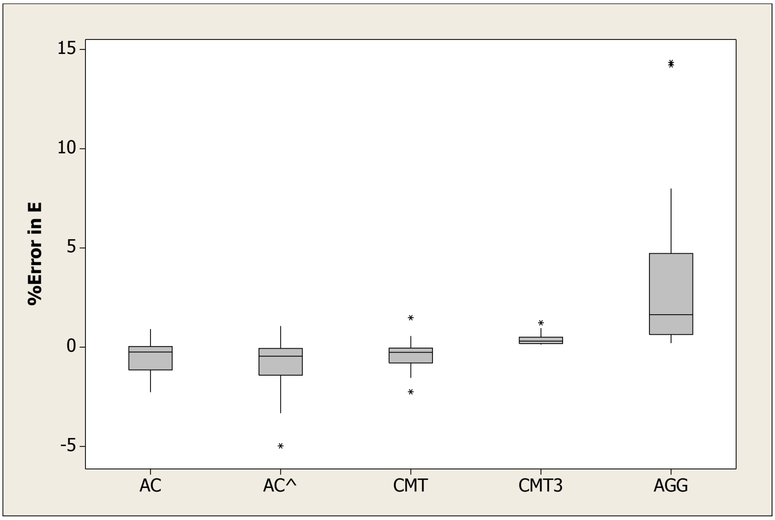


Figure 6.3: Boxplot for the percentage errors in calculating E with the proposed methods vs simulation.

Table 6.4: The average production rate E from applying different approximate methods

Case Number	Simulation LCL	Simulation Avg	Simulation UCL	AC	$\hat{AC}$	CMT	CMT3	AGG
1	0.6825	0.6826	0.6826	0.6879	0.6879	0.6830	0.6742	0.6650
2	0.6211	0.6212	0.6213	0.6203	0.6203	0.6198	0.6154	0.6110
3	0.8448	0.8449	0.8450	0.8449	0.8449	0.8447	0.8425	0.8405
4	0.8315	0.8317	0.8319	0.8313	0.8313	0.8313	0.8306	0.8300
5	0.7308	0.7308	0.7309	-	0.7314	0.7331	0.7279	0.7237
6	0.6552	0.6553	0.6554	0.6583	0.6583	0.6556	0.6490	0.6457
7	0.8652	0.8652	0.8653	0.8671	0.8671	0.8660	0.8626	0.8612
8	0.8393	0.8395	0.8396	0.8397	0.8397	0.8396	0.8380	0.8374
9	0.8112	0.8113	0.8113	-	0.8149	0.8130	0.8096	0.8082
10	0.7426	0.7427	0.7428	0.7521	0.7521	0.7460	0.7381	0.7310
11	0.9098	0.9099	0.9099	0.9140	0.9140	0.9116	0.9082	0.9053
12	0.8681	0.8683	0.8685	0.8703	0.8703	0.8695	0.8657	0.8621
13	0.8153	0.8154	0.8154	-	0.8179	0.8168	0.8143	0.8131
14	0.7507	0.7508	0.7509	0.7605	0.7597	0.7542	0.7474	0.7414
15	0.9134	0.9134	0.9135	-	0.9170	0.9151	0.9123	0.9098
16	0.8721	0.8723	0.8724	0.8746	0.8746	0.8736	0.8697	0.8664
17	0.6029	0.6029	0.6030	-	0.6194	0.6097	-	0.5164
18	0.4775	0.4776	0.4777	0.4763	0.4763	0.4705	-	0.4097
19	0.7365	0.7366	0.7367	0.7382	0.7382	0.7359	-	0.6936
20	0.6827	0.6830	0.6832	0.6813	0.6813	0.6809	-	0.6656
21	0.7275	0.7275	0.7276	-	0.7319	0.7315	-	0.7136
22	0.6217	0.6218	0.6219	-	0.6425	0.6294	-	0.5752
23	0.8467	0.8468	0.8469	-	0.8447	0.8526	-	0.8215
24	0.7661	0.7663	0.7665	0.7789	0.7789	0.7723	-	0.7277
25	0.7974	0.7974	0.7975	-	0.8081	0.8040	-	0.7844
26	0.6929	0.6930	0.6930	-	0.7274	0.7085	-	0.6376
27	0.8870	0.8871	0.8872	-	0.9035	0.8957	-	0.8608
28	0.7979	0.7981	0.7982	0.8156	0.8156	0.8077	-	0.7485
29	0.8116	0.8116	0.8116	-	0.8165	0.8151	-	0.8053
30	0.7297	0.7298	0.7299	-	0.7531	0.7411	-	0.6961
31	0.9046	0.9047	0.9047	-	0.9113	0.9101	-	0.8906
32	0.8291	0.8292	0.8294	0.8480	0.8480	0.8393	-	0.7869

Table 6.5: The output asymptotic variance rate V from applying different approximate methods

Case Number	Simulation LCL	Simulation Avg	Simulation UCL	AC	$\hat{AC}$	CMT	CMT3	AGG
1	0.3794	0.4426	0.5230	0.4187	0.4187	0.4360	0.4232	0.4115
2	1.4903	1.7385	2.0547	1.8343	1.8343	1.8678	1.8489	1.8346
3	0.8085	0.9431	1.1146	1.0770	1.0770	1.0894	1.0942	1.0999
4	4.2811	4.9941	5.9023	4.6299	4.6299	4.6388	4.6489	4.6570
5	0.3328	0.3883	0.4589	-	0.3432	0.4103	0.3771	0.3686
6	1.4081	1.6427	1.9414	1.7744	1.7744	1.8390	1.7838	1.7731
7	0.8157	0.9515	1.1246	0.9390	0.9390	0.9766	0.9643	0.9681
8	3.7205	4.3401	5.1294	4.4441	4.4441	4.4772	4.4870	4.4946
9	0.2209	0.2577	0.3046	-	0.2960	0.3560	0.2792	0.2393
10	1.2111	1.4128	1.6697	1.4628	1.4628	1.6935	1.5600	1.4658
11	0.5041	0.5881	0.6950	0.5836	0.5836	0.6976	0.6387	0.6142
12	3.8926	4.5408	5.3666	3.6763	3.6763	3.8882	3.8596	3.8779
13	0.2073	0.2418	0.2858	-	0.2888	0.4118	0.2712	0.2325
14	1.1932	1.3919	1.6450	1.4317	1.4257	1.6596	1.5061	1.4132
15	0.4427	0.5164	0.6104	-	0.5533	0.7112	0.6004	0.5743
16	2.8235	3.2938	3.8928	3.5580	3.5580	3.7999	3.7400	3.7537
17	0.2118	0.2470	0.2920	-	0.3506	0.4054	-	0.2853
18	0.9170	1.0697	1.2642	1.3382	1.3382	1.4395	-	1.1390
19	0.9564	1.1156	1.3185	1.3104	1.3104	1.3814	-	1.3840
20	4.8199	5.6226	6.6451	5.7505	5.7505	5.8468	-	5.8251
21	0.2923	0.3409	0.4030	-	0.2768	0.4816	-	0.3144
22	1.0738	1.2526	1.4804	-	0.8791	1.8194	-	1.3347
23	0.7955	0.9280	1.0968	-	1.1986	1.0802	-	0.9710
24	3.7365	4.3587	5.1514	4.8172	4.8172	5.3420	-	5.3012
25	0.1442	0.1682	0.1988	-	0.2938	0.3694	-	0.1749
26	0.8917	1.0402	1.2294	-	1.3606	1.7688	-	1.1989
27	0.4682	0.5461	0.6454	-	0.5883	0.8315	-	0.7021
28	3.1736	3.7022	4.3754	4.1792	4.1792	4.9365	-	5.0579
29	0.1590	0.1855	0.2192	-	0.3008	0.4259	-	0.1706
30	0.8041	0.9380	1.1086	-	1.1556	1.7441	-	1.1054
31	0.3682	0.4295	0.5076	-	0.4888	0.7640	-	0.5272
32	2.8027	3.2694	3.8640	3.6455	3.6455	4.4843	-	4.4715

Table 6.6: The evaluation time T_{ev} (in seconds) for applying different approximate methods

Case Number	Simulation	AC	$\hat{AC}$	CMT	CMT3	AGG
1	1957	1	1	9	153	0.01
2	1919	1	1	11	154	0.01
3	1998	1	1	9	157	0.02
4	1992	1	1	14	159	0.03
5	1985	1	5	10	199	0.02
6	1937	1	1	14	222	0.02
7	2008	1	1	10	306	0.02
8	1991	1	1	10	317	0.04
9	2033	3	3	19	12516	0.04
10	1990	4	3	19	12366	0.04
11	2029	3	3	19	12706	0.05
12	2005	2	2	15	12714	0.06
13	2027	3	4	6	28524	0.05
14	1991	5	4	12	27759	0.04
15	2031	1	4	6	25964	0.06
16	2011	2	2	10	25200	0.07
17	5036	2	6	116	-	0.03
18	4809	3	2	163	-	0.03
19	5094	2	2	105	-	0.04
20	5009	2	2	157	-	0.07
21	5225	7	21	97	-	0.08
22	5065	6	21	160	-	0.08
23	5264	3	21	110	-	0.09
24	5149	7	6	183	-	0.12
25	5324	13	13	154	-	0.12
26	5176	14	26	185	-	0.10
27	5323	17	23	106	-	0.13
28	5189	19	10	168	-	0.15
29	5351	25	51	179	-	0.20
30	5233	24	51	596	-	0.16
31	5342	27	52	238	-	0.21
32	5254	32	25	342	-	0.22

Table 6.7: A comparison between the accuracy in calculating E for the different approximate analytical methods used for the evaluation of the factorial design

Case Number	AC	$\hat{AC}$	CMT	CMT3	AGG
1	-0.776%	-0.776%	-0.064%	1.227%	2.57%
2	0.136%	0.136%	0.213%	0.929%	1.642%
3	-0.002%	-0.002%	0.017%	0.285%	0.523%
4	0.044%	0.044%	0.045%	0.131%	0.204%
5	-	-0.079%	-0.309%	0.405%	0.978%
6	-0.458%	-0.458%	-0.055%	0.959%	1.456%
7	-0.214%	-0.214%	-0.086%	0.304%	0.471%
8	-0.03%	-0.03%	-0.021%	0.177%	0.243%
9	-	-0.452%	-0.216%	0.209%	0.377%
10	-1.269%	-1.269%	-0.446%	0.618%	1.566%
11	-0.456%	-0.456%	-0.186%	0.181%	0.5%
12	-0.228%	-0.228%	-0.139%	0.305%	0.719%
13	-	-0.317%	-0.182%	0.127%	0.282%
14	-1.291%	-1.18%	-0.447%	0.462%	1.262%
15	-	-0.394%	-0.179%	0.129%	0.395%
16	-0.26%	-0.26%	-0.153%	0.29%	0.67%
17	-	-2.742%	-1.124%	-	14.343%
18	0.269%	0.269%	1.485%	-	14.218%
19	-0.216%	-0.216%	0.092%	-	5.842%
20	0.241%	0.241%	0.299%	-	2.547%
21	-	-0.608%	-0.546%	-	1.916%
22	-	-3.329%	-1.226%	-	7.492%
23	-	0.247%	-0.685%	-	2.983%
24	-1.646%	-1.646%	-0.789%	-	5.032%
25	-	-1.336%	-0.817%	-	1.631%
26	-	-4.972%	-2.249%	-	7.995%
27	-	-1.847%	-0.975%	-	2.958%
28	-2.195%	-2.195%	-1.213%	-	6.203%
29	-	-0.599%	-0.434%	-	0.774%
30	-	-3.193%	-1.548%	-	4.615%
31	-	-0.728%	-0.596%	-	1.552%
32	-2.27%	-2.27%	-1.211%	-	5.103%
Avg errors for applicable cases:	-0.59%	-0.964%	-0.429%	0.421%	3.096%
Avg abs errors for applicable cases:	0.911%	1.065%	0.564%	0.421%	3.096%

Table 6.8: A comparison between the accuracy in calculating V for the different approximate analytical methods used for the evaluation of the factorial design

Case Number	AC	$\hat{AC}$	CMT	CMT3	AGG
1	5.392%	5.392%	1.472%	4.367%	7.016%
2	-5.511%	-5.51%	-7.435%	-6.347%	-5.524%
3	-14.193%	-14.193%	-15.509%	-16.017%	-16.624%
4	7.292%	7.292%	7.113%	6.912%	6.749%
5	-	11.605%	-5.664%	2.878%	5.059%
6	-8.02%	-8.02%	-11.951%	-8.594%	-7.939%
7	1.312%	1.312%	-2.637%	-1.337%	-1.744%
8	-2.395%	-2.395%	-3.158%	-3.382%	-3.558%
9	-	-14.842%	-38.142%	-8.329%	7.161%
10	-3.539%	-3.539%	-19.87%	-10.42%	-3.748%
11	0.753%	0.753%	-18.632%	-8.61%	-4.438%
12	19.038%	19.038%	14.371%	15.001%	14.599%
13	-	-19.429%	-70.295%	-12.124%	3.841%
14	-2.86%	-2.43%	-19.234%	-8.203%	-1.53%
15	-	-7.13%	-37.715%	-16.257%	-11.197%
16	-8.021%	-8.021%	-15.365%	-13.547%	-13.963%
17	-	-41.9%	-64.087%	-	-15.48%
18	-25.097%	-25.097%	-34.574%	-	-6.476%
19	-17.454%	-17.454%	-23.816%	-	-24.051%
20	-2.273%	-2.274%	-3.987%	-	-3.6%
21	-	18.828%	-41.264%	-	7.794%
22	-	29.816%	-45.252%	-	-6.558%
23	-	-29.154%	-16.396%	-	-4.633%
24	-10.519%	-10.519%	-22.557%	-	-21.623%
25	-	-74.635%	-119.545%	-	-3.958%
26	-	-30.794%	-70.034%	-	-15.251%
27	-	-7.718%	-52.254%	-	-28.564%
28	-12.886%	-12.886%	-33.34%	-	-36.62%
29	-	-62.154%	-129.587%	-	8.021%
30	-	-23.196%	-85.94%	-	-17.843%
31	-	-13.822%	-77.895%	-	-22.744%
32	-11.502%	-11.502%	-37.159%	-	-36.767%
Avg errors for applicable cases:	-5.026%	-11.08%	-34.385%	-5.25%	-7.943%
Avg abs errors for applicable cases:	8.781%	16.958%	35.82%	8.895%	11.709%

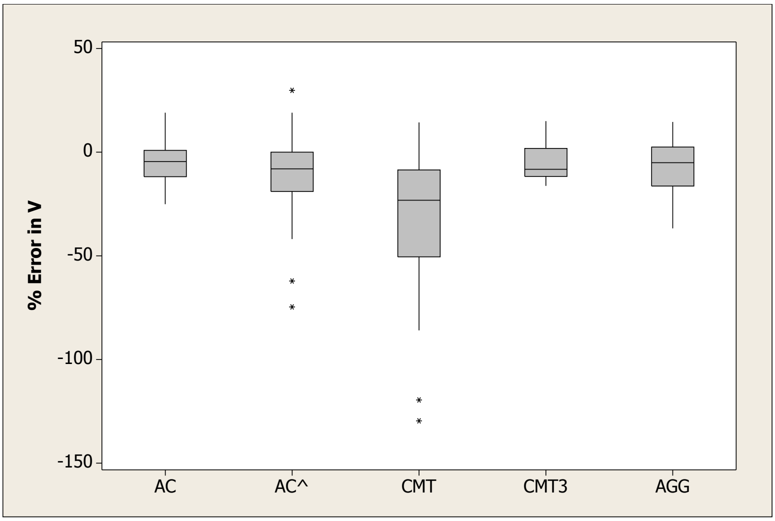


Figure 6.4: Boxplot of the percentage errors in calculating V with the proposed methods vs simulation.

Table 6.9: Significant factors for each of the studies approximate methods

Method	Significant factors
AC	Line Length, Buffer Capacity and isolated machine's efficiency
$\hat{AC}$	Line Length, Buffer Capacity and isolated machine's efficiency
CMT	Line Length, Buffer Capacity , (Buffer Capacity+Increment factor) , (Buffer Capacity+ isolated machine's efficiency)
CMT3	isolated machine's efficiency
AGG	Line Length

of additional errors compared with the AC method. The Aggregation method somehow trades off the accuracy in estimating E with the accuracy in estimating V.

In terms of speed, the Aggregation method is the fasted among all the methods due to two reasons: the low number of states it works with and the fact that it does not iterate. Next comes the other decomposition methods CMT, AC and $\hat{AC}$, and finally the CMT3 which is feasible to apply in limited cases.

6.2.2 Factors Affecting the Accuracy of the Proposed Methods

For investigating the effect of the line and machine parameters on the accuracy of the methods during the estimation of V, the Lenth's method was used for unreplicated experiments. Figures [6.5, 6.6, 6.7, 6.8, 6.9] show the application of the Lenth's method for the five proposed approaches finding the factors that are significant. These factors are summarized in Table 6.9.

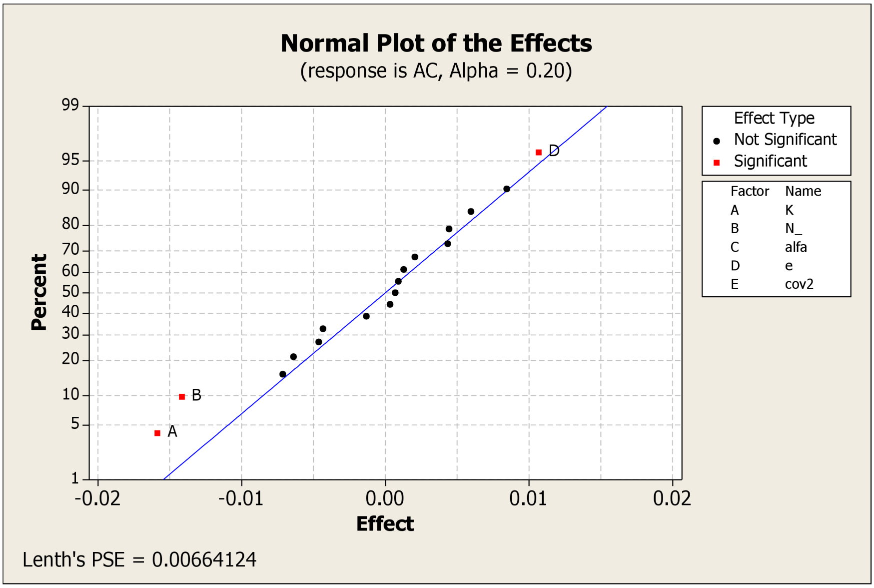


Figure 6.5: The application of Lenth's method on the AC method

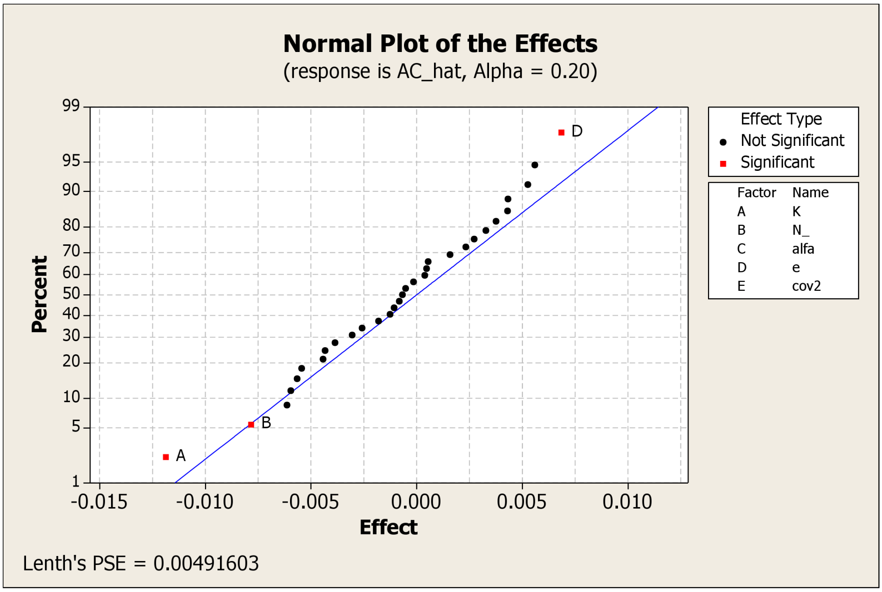


Figure 6.6: The application of Lenth's method on the $\hat{AC}$ method

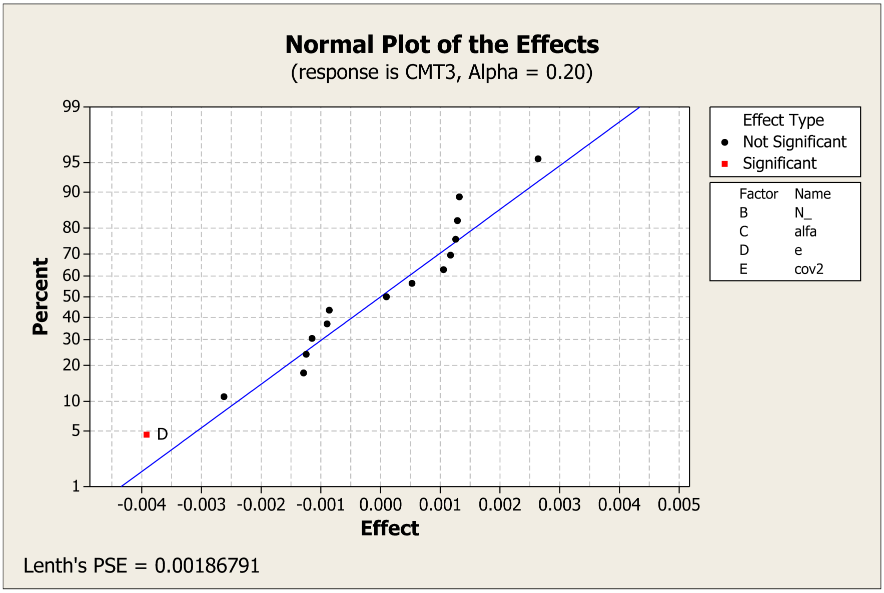


Figure 6.7: The application of Lenth's method on the CMT method

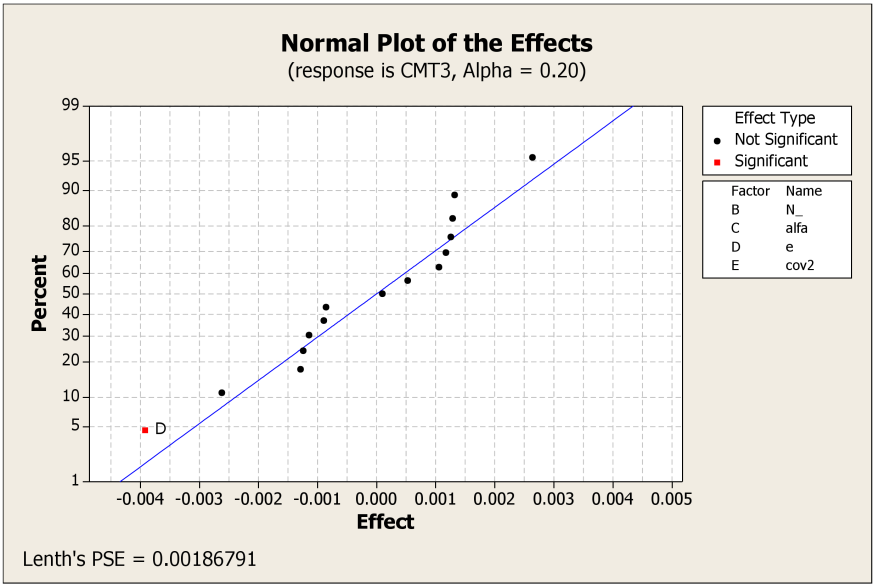


Figure 6.8: The application of Lenth's method on the CMT3 method

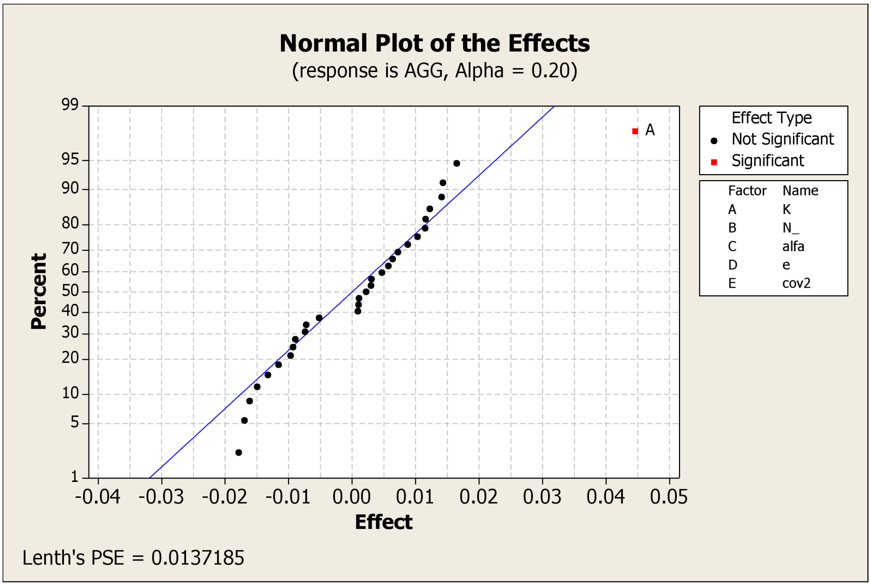


Figure 6.9: The application of Lenth's method on the Aggregation method

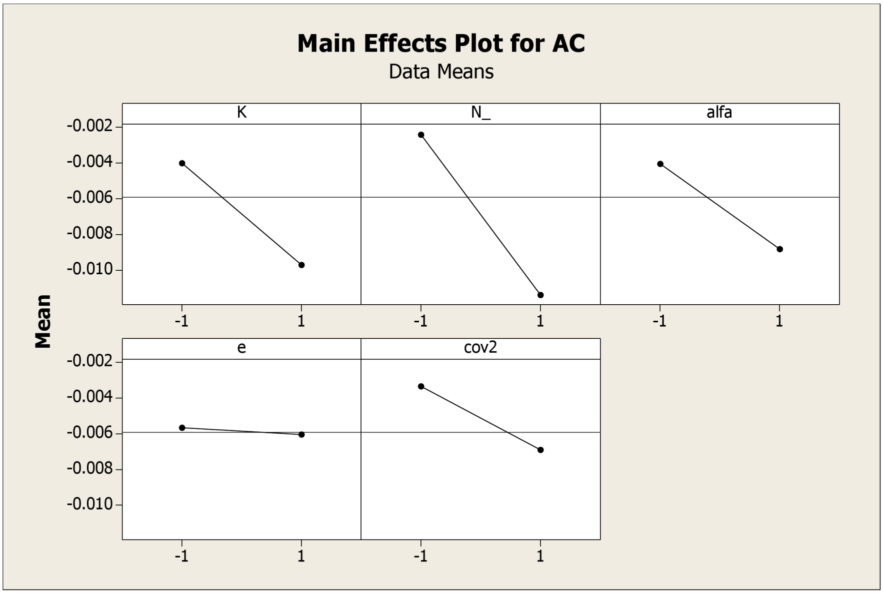


Figure 6.10: Main effects plot for AC method

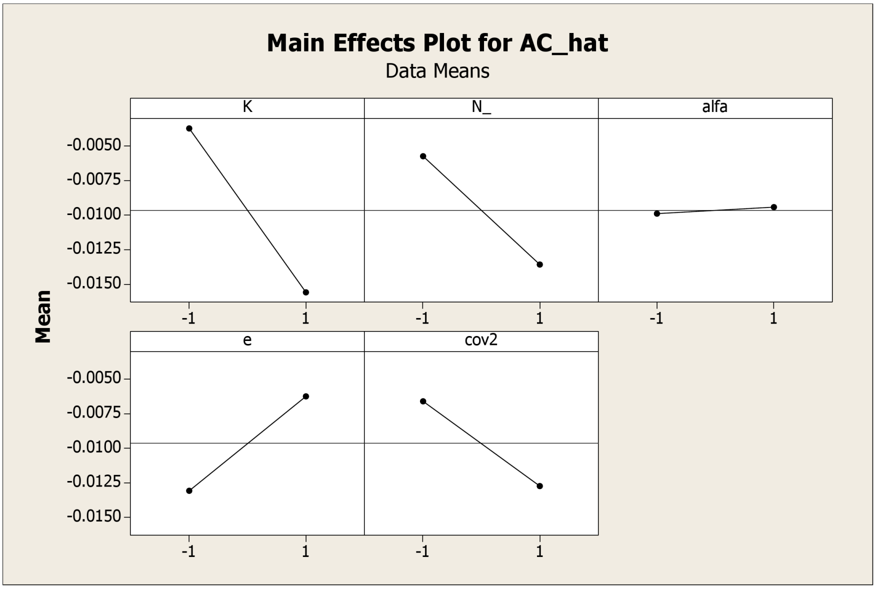


Figure 6.11: Main effects plot for $\hat{AC}$ method

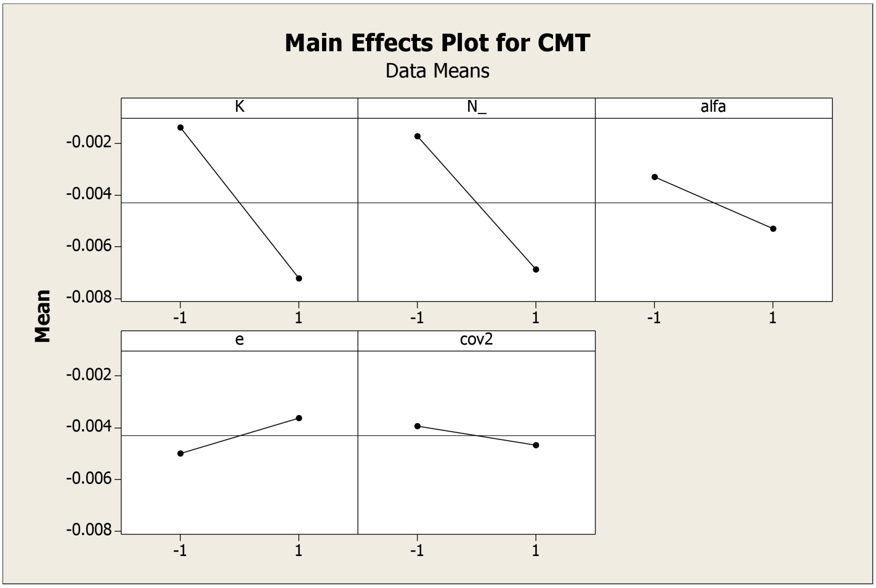


Figure 6.12: Main effects plot for CMT method

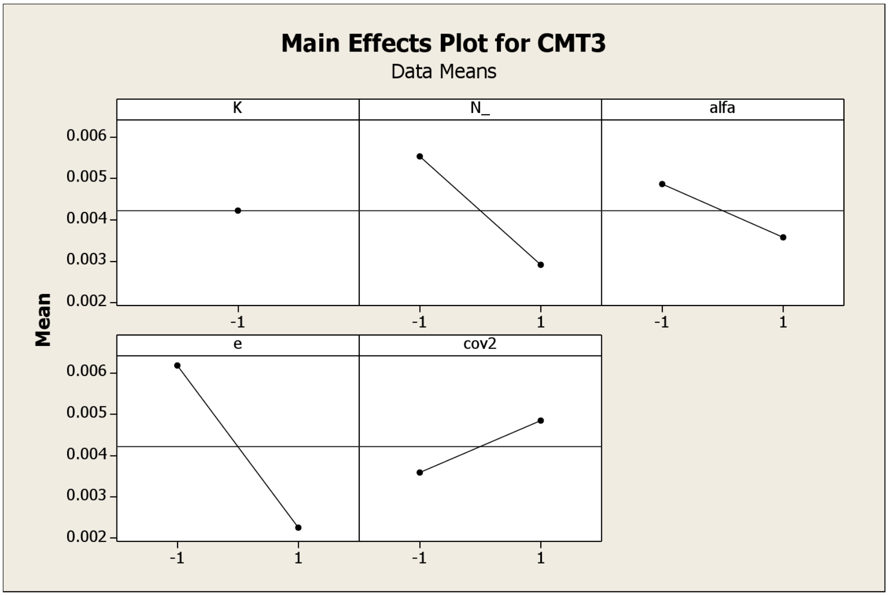


Figure 6.13: Main effects plot for CMT3 method

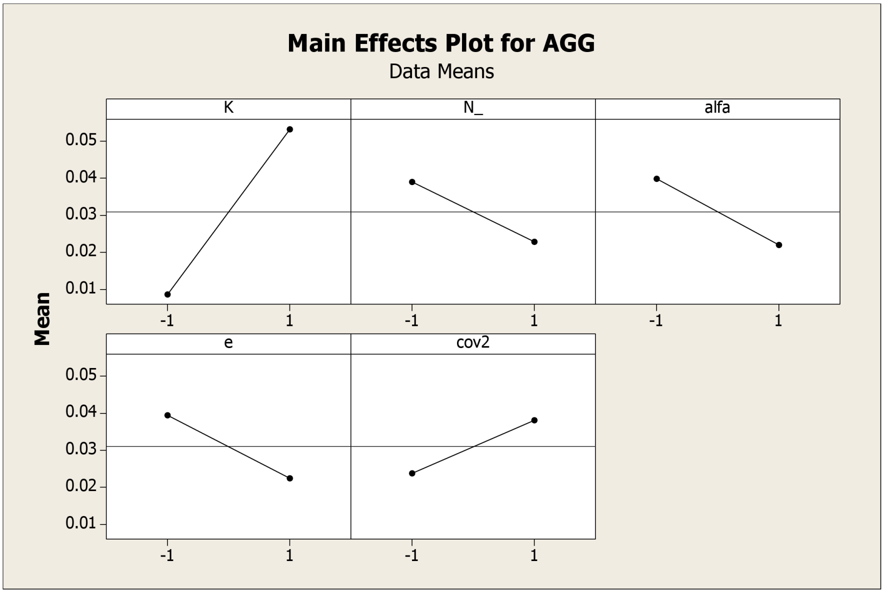


Figure 6.14: Main effects plot for Aggregation method

```
Factorial Fit: AC versus K; N; e

Estimated Effects and Coefficients for AC (coded units)

Term            Effect        Coef    SE Coef       T       P

Constant                 -0.008520  0.001810   -4.71   0.000

K            -0.007046   -0.003523  0.001669   -2.11   0.053

N            -0.009698   -0.004849  0.001568   -3.09   0.008

e             0.002190    0.001095  0.001669    0.66   0.522


S = 0.00642876   PRESS = 0.00104832

R-Sq = 47.78%    R-Sq(pred) = 5.40%   R-Sq(adj) = 36.59%
```

ANOVA results for the AC method

It can be seen from the ANOVA analysis for the AC method that it is neither affected by the line length nor the efficiency in isolation of the machines.

ANOVA results for the $\hat{AC}$ method

It can be seen from the output of the ANOVA test of MINITAB that the $\hat{AC}$ is affected by the factors Line length (K) and buffer capacity (N) and efficiency in isolation (e). This is due to the approximation used when the AC method does not work.

ANOVA results for the CMT method

The CMT method as well is affected by the line length and buffer capacities. The average accuracy for estimating V for the method decreases from -15.1% to -53.6% when changing the line length from 4 to 10, and drops from -18.1% to -50.6% when changing the buffer capacities from 3 to 15.

```
Factorial Fit: AC_hat versus K; N; e

Estimated Effects and Coefficients for AC_hat (coded units)


Term           Effect       Coef   SE Coef      T      P
Constant                -0.009647  0.001764  -5.47  0.000
K           -0.011869  -0.005934  0.001764  -3.36  0.002
N           -0.007837  -0.003919  0.001764  -2.22  0.035
e            0.006845   0.003423  0.001764   1.94  0.062


S = 0.00997981   PRESS = 0.00364239
R-Sq = 41.68%    R-Sq(pred) = 23.83%   R-Sq(adj) = 35.43%
```

ANOVA results for the CMT3 method

The CMT3 method is only affected significantly by factor e, when e changes from 0.85 to 0.95 the accuracy of the method increases from -5.87% to -4.65%.

ANOVA results for the AGG method

Finally, the Aggregation method's accuracy for calculating V also drops from -1.62% to -14.3% by changing the line length from 4 to 10 machines.

General Results

The residuals of the ANOVA test have been checked to be normally distributed and not correlated. As it can be noticed, while considering a $\alpha_{\text{family}} = 0.05$ only some of the factors introduced were significant (see the reported results from MINITAB™ 15).

The ANOVA analysis shows that $\hat{AC}$, CMT and AGG are affected by the line length i.e. when the number of machines change from 4 to 10. CMT is affected the most as its average accuracy drops from -15.1% to -53.6% on the contrary to the other methods, which are affected much less.

```
Factorial Fit: CMT versus K; N; alfa; e

Estimated Effects and Coefficients for CMT (coded units)
```

Term	Effect	Coef	SE Coef	T	P
Constant		-0.004298	0.000956	-4.49	0.000
K	-0.005830	-0.002915	0.000956	-3.05	0.005
N	-0.005153	-0.002577	0.000956	-2.69	0.012
alfa	-0.001996	-0.000998	0.000956	-1.04	0.307
e	0.001367	0.000683	0.000956	0.71	0.482
N *alfa	0.003861	0.001930	0.000956	2.02	0.054
N *e	0.000742	0.000371	0.000956	0.39	0.701

```
S = 0.00541019   PRESS = 0.00119891

R-Sq = 47.23%    R-Sq(pred) = 13.53%   R-Sq(adj) = 34.56%
```

6.3 Output Asymptotic Variance Rate Behavior in Multi-Stage Production Lines

In Chapter 4 a discussion was given on the effect of the buffer capacity on a two machine system using the exact method proposed. In this section, the approximate methods will be used to study the system behavior in a three machine line. The goal of this section is to identify in what cases the modification of the buffer capacity is beneficial i.e. reducing the output variability.

6.3.1 Three Machine Line - The Effect of Buffer Capacities

The system consists of three machines with efficiencies in isolation e equal to [0.92, 0.88, 0.82] and two variable capacity buffer spaces N_1 and N_2. What is the effect of the buffer capacities on the output performances, when the machines have different values of v? To

```
Factorial Fit: CMT3 versus e

Estimated Effects and Coefficients for CMT3 (coded units)


Term           Effect        Coef    SE Coef       T      P

Constant                 0.004217   0.000703   6.00   0.000

e          -0.003919   -0.001960   0.000703   -2.79   0.014


S = 0.00281015    PRESS = 0.000144401

R-Sq = 35.72%    R-Sq(pred) = 16.04%   R-Sq(adj) = 31.13%
```

investigate the behavior, three different cases have been evaluated using Aggregation method then studied and reported.

Case 1

Assume the three machine have $[e, v]$ as following: machine M_1: [0.92, 8] machine M_2 [0.88, 5] and machine M_3 [0.82, 2]. In this scenario the bottleneck machine is the machine with the lowest v i.e. M_3. It is known in the case of a two machine line V approaches v of the bottleneck machine, so we expect adding buffers would always make V decrease, and in fact this is the case. Figure 6.15 shows that by increasing the capacity of the buffers N_1 and N_2, V decreases. The only important thing is to notice that the effect of increasing N_1 is not the same as the effect of increasing N_2. The reason is because N_2 is the buffer before the bottleneck machine and when decoupling it with buffers V approaches v of M_3. Systems with such configuration will only decrease the V by increasing buffers capacities on all buffers.

Case 2

This case is the opposite of case 1 in terms of v, the machine with highest e has also the highest v. Values of $[e, v]$ are as following: machine M_1:[0.92, 2] machine M_2:[0.88, 5] and machine M_3:[0.82, 8]. Figure 6.16 shows that by increasing the capacity of N_2, V decreases. While

```
Factorial Fit: AGG versus K

Estimated Effects and Coefficients for AGG (coded units)


Term        Effect     Coef    SE Coef     T      P

Constant             0.03096  0.005199  5.95  0.000

K          0.04459  0.02230  0.005199  4.29  0.000


S = 0.0294127   PRESS = 0.0295290

R-Sq = 38.00%   R-Sq(pred) = 29.46%   R-Sq(adj) = 35.93%
```

increasing N_1 has a negative effect on V, this is very important to notice when considering the allocation of buffers in the system.

Case 3

The machine with highest v is the machine that is 2nd bottleneck machine. Values of $[e, v]$ are as following: machine M_1:[0.92, 2] machine M_2:[0.88, 8] and machine M_3:[0.82, 5]. This case also is not intuitive, the system's V is minimum when N_2 is large and N_1 is neither small nor large is capacity, Figure 6.17 shows such relationship.

6.3.2 Three Machine Line - Reversibility Property of the Asymptotic Variance Rate

The works of Carrascosa [11], Hendricks [42] only studied the reversibility property of the asymptotic behavior of V in two machine line with one intermediate buffer. Here, we numerically tested longer lines, showing that the difference between the output variability in the forward system (forward flow of parts) and its reverse are very close. Table 6.10 shows that the differences between the lines are almost negligible.

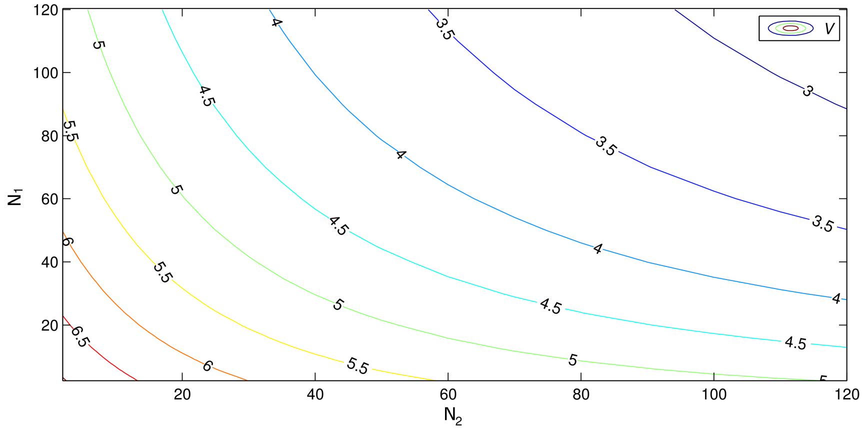


Figure 6.15: Case1 - The effect of changing the buffer capacity of buffers N_1 and N_2 on V

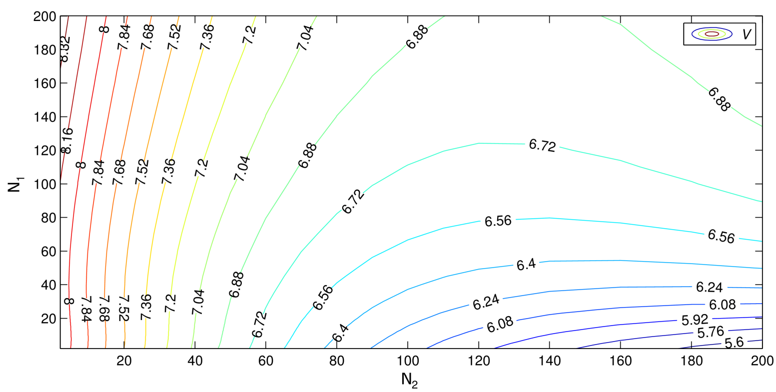


Figure 6.16: Case2 - The effect of changing the buffer capacity of buffers N_1 and N_2 on V

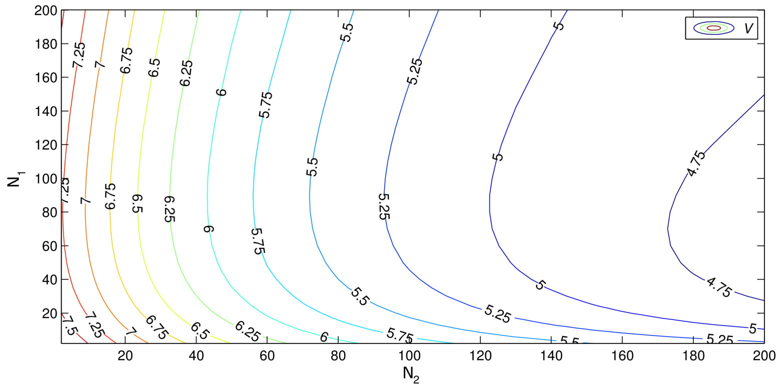


Figure 6.17: Case3 - The effect of changing the buffer capacity of buffers N_1 and N_2 on V

Table 6.10: Comparing V and E of a set of three machine systems and their reverse

Line Parameters								Forward System				Backward System			
e_1	e_2	e_3	v_1	v_2	v_3	N_1	N_2	V	E	$\bar{n_1}$	$\bar{n_2}$	V	E	$\bar{n_1}$	$\bar{n_2}$
0.92	0.82	0.85	4	1.6	2.5	4	3	3.6059	0.684	2.8316	1.2431	3.5964	0.6842	1.6737	1.001
0.644	0.574	0.595	4	1.6	2.5	5	2	1.3385	0.358	3.0294	0.8298	1.3053	0.3634	1.0257	1.3955
0.92	0.82	0.85	2.8	1.12	1.75	2	5	2.5326	0.6967	1.1015	1.8954	2.5665	0.6955	2.927	0.8157
0.77	0.89	0.85	6	0.8	1.25	3	3	4.0017	0.6387	1.4137	1.2349	3.9941	0.6385	1.4081	1.3323
0.77	0.89	0.85	8.4	1.12	1.75	3	3	5.5994	0.6352	1.4185	1.2324	5.5916	0.6349	1.4098	1.3249
0.77	0.89	0.85	18	2.4	3.75	6	3	11.9529	0.6338	3.0132	1.2438	11.8552	0.6352	1.4122	2.2792
0.77	0.89	0.85	6	0.8	1.25	6	3	3.9792	0.6477	2.8724	1.2754	3.8868	0.6522	1.4081	2.3797

6.4 Conclusions

From this chapter, a comparison between different methods have been performed, showing factors that significantly affect their accuracy. The AC method is the most accurate for calculating V and E, however, it does not always work. While its alternative $\hat{AC}$ which always work has a tradeoff behavior as it has additional errors in calculating E and V.

The CMT method is not accurate in calculating V, the factorial design shows that errors can be very high. However, its alternative CMT3 works better with the cost of extensive computational time. Finally, the Aggregation method is an acceptable method for calculating V in terms of accuracy, however, it gives higher errors in calculating E.

The effect of changing two buffer capacities in a three machine line has been investigated, showing that the V surface can have continuous non-monotonic structure, this will be helpful for the optimal buffer allocation problem that considers the the asymptotic variance rate.

Chapter 7

Optimization of Production Systems Considering the Output Variability

In the previous chapter, the analysis of the system behavior has been performed, showing a strong relationship between the output variability and the buffer capacities in the line. The analysis highlighted the need for tools and methods able to support the production system designer/manager while configuring the system. These tools and methods should be able to take into consideration the output variability of the system. In this chapter, the problem of optimally designing a production system in order to minimize the output variability is addressed proposing different methods to model the problem.

The work will introduce a new method for the Buffer Allocation Problem (BAP) in transfer lines, taking into account the output variability. This method is an extension to methods proposed for first order performance measures (i.e. The gradient method) such as the method proposed by Gershwin and Goldis [32].

The performance measures optimized with such approach are: The asymptotic variance rate V, and the service level (SL). The optimization procedure will be tested on three machine line systems, and will be used to optimize one real system presented in Chapter 8.

7.1 Buffer Allocation Problem

In the literature, different allocation problems and formulations have been defined. These formulations are of interest for industrial applications. Buffer allocation problems can be, from one point of view, either constrained or unconstrained problems [[32], [75]].

7.1.1 Unconstrained Buffer Allocation Problem

The aim in this approach is to find the vector N of buffer sizes that minimizes the output variability of the system. This approach is the simplest in terms of formulation, and can be represented by using the formalism of linear programming as:

$$\text{minimize} \qquad V(N_1, N_2, \ldots N_{K-1}) \tag{7.1}$$

Two-machine line cases represented in chapter 4 and studies by [11, 82] and Colledani et al. [19] show that the output asymptotic variance rate V as a function of buffer sizes, can have a convex form, thus it would make sense to use such formulation to find a minimum for the function.

7.1.2 Constrained Buffer Allocation Problem

The constrained BAP can be either primal or dual problems. In the primal problem the aim is to find out the buffer space configuration $(N_1, \ldots, N_{K-1})$ in a linear multi-stage line, that minimizes the total buffer space in order to achieve a target production rate E^{obj}, a certain output variability requirement V^{obj} or a service level requirement SL^{obj}. The formalization of the primal constrained problem, can be in the form of the following equations:

$$\begin{aligned} \text{minimize} \qquad & N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i \\ \text{subject to} \qquad & E(N_1, \ldots, N_{K-1}) \geq E^{obj} \\ & N_i \geq N^{MIN} = 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned} \tag{7.2}$$

or

$$\begin{aligned} &\text{minimize} && N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i \\ &\text{subject to} && V(N_1, \ldots, N_{K-1}) \leq V^{obj} \\ & && N_i \geq N^{MIN} = 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned} \tag{7.3}$$

or

$$\begin{aligned} &\text{minimize} && N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i \\ &\text{subject to} && SL(N_1, \ldots, N_{K-1}) \geq \text{SL}^{obj} \\ & && N_i \geq N^{MIN} = 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned} \tag{7.4}$$

or even a mix of performance measures requirements, such as:

$$\text{minimize} \quad N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i$$

$$\text{subject to} \quad SL(N_1, \ldots, N_{K-1}) \geq \text{SL}^{obj} \tag{7.5}$$

$$V(N_1, \ldots, N_{K-1}) \leq V^{obj} \tag{7.6}$$

$$N_i \geq N^{MIN} = 3, \text{i} = 1, \ldots, \text{K-1}$$

The Dual BAP problem has the goal to optimize a certain performance measure such as, to maximize the production rate and the SL, or to minimize the output variability. The Dual BAP is subject to the total buffer space constraint (N^{total}). The formalization of the constrained Dual BAP is given as:

$$\begin{aligned} &\text{maximize} && E(N_1, \ldots, N_{K-1}) \\ &\text{subject to} && N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i \\ &\text{subject to} && N_i \geq N^{MIN} = 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned} \tag{7.7}$$

Using the same formalism, other performance measures can be optimized as:

$$\begin{aligned} \text{minimize} \quad & V(N_1, \ldots, N_{K-1}) \\ \text{subject to} \quad & N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i \\ & N_i \geq N^{MIN} = 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned} \tag{7.8}$$

or

$$\begin{aligned} \text{Maximize} \quad & SL(N_1, \ldots, N_{K-1}) \\ \text{subject to} \quad & N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i \\ & N_i \geq N^{MIN} = 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned} \tag{7.9}$$

7.1.3 Solution of the Buffer Allocation Problem

The gradient method has been used for solving the BAP considering the average throughput only. The goal of the method is to find with a certain buffer capacity the maximum throughput. The gradient is estimated by evaluating the performance of alternative buffer allocations which are in the neighbourhood of the allocation under analysis. Then, based on the direction of improvement the new buffer allocation is assigned. Details for the application on the average throughput of the system can be found in Gershwin and Schor [37].

Gradient based methods were used extensively in calculating the Primal and Dual BAP that consider the average throughput. The throughput curve (as a function of the buffer capacities) has a particular and known behavior. The work of Gershwin and Schor [37] show that the throughput curve has the following properties:

- Monotonicity: The throughput function of the system $E = f(N_1, \ldots, N_K)$ increases monotonically in each N_i;
- Concavity: The function $E = f(N_1, \ldots, N_K)$ is a concave function.

As a sequence, the throughput function has an upper limit that is given by the minimal efficiency in isolation mode, among the machines in the line. More details about the primal and dual BAP can be found for example in Papadopoulos et al. [70].

The problem with the output asymptotic variance rate surfaces is that they are usually not monotonic (see Section 4.2), this could create problems in finding the optimal buffer allocation. To solve this issue the Non-linear Conjugate Gradient method (CG) of Fletcher-Reeves-Polak-Ribière (FR-PR) is used. [for more details see [66]]. These optimization problems are very sensitive to the step size at iteration k (α_k), in fact the step size should satisfy the Wolfe conditions, which are:

$$\boldsymbol{V}(\boldsymbol{N}_k + \alpha_k \boldsymbol{p}_k) \leq \boldsymbol{V}(\boldsymbol{N}_k) + c_1 \alpha_k \nabla \boldsymbol{V}_k' \boldsymbol{p}_k \tag{7.10}$$

$$|\nabla V(\boldsymbol{N}_k + \alpha_k \boldsymbol{p}_k)' \boldsymbol{p}_k| \leq -c_2 \nabla \boldsymbol{V}_k' \boldsymbol{p}_k \tag{7.11}$$

where ($'$) indicates the matrix transpose operation, and $0 < c_1 < c_2 < 0.5$, and ∇V_k is the gradient of V at iteration k, and $\boldsymbol{p}_k$ is the vector of directions of improvement at iteration k. Furthermore, for any α_k that satisfies these conditions, all the directions $\boldsymbol{p}_k$ are decent directions for V (see [66] for the proof).

7.1.4 The Conjugate Gradient Algorithm for the Output Variability Optimization

To explain the procedure of applying the Conjugate Gradient (CG) algorithm, let's consider a system of a m-machine line and $m-1$ intermediate buffers. The goal of the approach is to find out an optimal buffer allocation of $\boldsymbol{N}$ that gives the minimum V.

1. *Initialization:* Set iteration number $k = 1$ and set $\boldsymbol{N}^{\text{initial}} = \boldsymbol{N}^{MIN}$ as the vector of initial buffer capacities. Evaluate the system with $\boldsymbol{N}^{\text{initial}}$ and store the output as a vector $\boldsymbol{V}_{k-1}$ i.e $\boldsymbol{V}_0$ with V in each element of the vector. It should be mentioned, if we are trying to optimize a constrained BAP, the solution of the unconstrained problem can be a good starting point [see [75]].

2. *Find out the new $\boldsymbol{V}$ vector for the next step:* Evaluate the system finding $\boldsymbol{V}_k$ as:

$$\boldsymbol{V}_k = [V_{k,\{1\}}, V_{k,\{2\}}, \ldots, V_{k,\{m-1\}}] \tag{7.12}$$

and $V_{k,\{i\}}$ is:

$$V_{k,\{i\}} = V(\boldsymbol{N}_{k-1} + [0, 0, 0 \ldots \alpha_k \ldots] \;\; \alpha_k \text{ added to buffer } i \tag{7.13}$$

3. *Calculating the gradient* $\nabla \boldsymbol{V}_k$*:*

$$\nabla \boldsymbol{V}_k = \frac{\boldsymbol{V}_k - \boldsymbol{V}_{k-1}}{||\boldsymbol{V}_k - \boldsymbol{V}_{k-1}||} \tag{7.14}$$

Negative gradient means $\boldsymbol{V}$ is decreasing, while positive $\boldsymbol{V}$ means that it is increasing.

4. Find out The Improvement Direction: The improvement direction $\boldsymbol{p}_k$ is calculated as:

$$\boldsymbol{p}_k = \frac{-\nabla \boldsymbol{V}_k}{||\nabla V_k||} \tag{7.15}$$

where $||\cdot||$ is the norm of the vector.

5. *Check the Wolfe conditions:* The Wolfe conditions represented in equations (7.10) and (7.11) should be satisfied. If α_k does not satisfy these equations then change it with one that does. Otherwise, if the conditions are not satisfied by any α_k, then a better solution does not exist compared with the current one.

6. *Calculate the new buffer capacities:* The new vector of buffer capacities $\boldsymbol{N}_{k+1}$ can be calculated as

$$\boldsymbol{N}_{k+1} = \boldsymbol{N}_k + \lceil \alpha_k \cdot \boldsymbol{p}_k \rceil \tag{7.16}$$

where $\lceil\cdot\rceil$ is the rounding up to the next integer operation.

7. *Evaluate the new* $\boldsymbol{V}_{k+1}$ *vector and gradients:* Steps 2 and 3 are used to calculate $\boldsymbol{V}_{k+1}$ and $\nabla \boldsymbol{V}_{k+1}$.

8. *Calculate* β_{k+1} *as:*

$$\beta_{k+1} = \begin{cases} -\beta_{k+1}^{FR} & \text{if } \beta_{k+1}^{PR} < -\beta_{k+1}^{FR} \\ \beta_{k+1}^{PR} & \text{if } |\beta_{k+1}^{PR}| \le \beta_{k+1}^{FR} \\ \beta_{k+1}^{FR} & \text{if } \beta_{k+1}^{PR} > \beta_{k+1}^{FR} \end{cases} \tag{7.17}$$

where β^{FR} is calculated by the Fletcher-Reeves (FR) method as:

$$\beta_{k+1}^{FR} = \frac{\nabla \boldsymbol{V}_{k+1}' \nabla \boldsymbol{V}_{k+1}}{\nabla \boldsymbol{V}_k' \nabla \boldsymbol{V}_k} \tag{7.18}$$

and β^{PR} is calculated by the Polak-Ribière (PR) method as:

$$\beta_{k+1}^{PR} = \frac{\nabla \boldsymbol{V}_{k+1}' (\nabla \boldsymbol{V}_{k+1} - \nabla \boldsymbol{V}_k)}{||\nabla \boldsymbol{V}_k||^2} \tag{7.19}$$

9. *Find the New direction as:*

$$p_{k+1} = \nabla \boldsymbol{V}_{k+1} + \beta_{k+1} \cdot \boldsymbol{p}_k \tag{7.20}$$

and set $k = k + 1$;

10. *Repetition :* Steps 2-9 are repeated until $\nabla \boldsymbol{V}_{k+1} = 0$.

7.1.5 Unconstrained Optimization

To test the proposed approach, a system composed of three machines and two intermediate buffers was used. Table 7.1 shows the parameters of the studied machines. The AGG method was used to evaluate the line.

Table 7.1: The case studied for the unconstrained optimization

Machine	Failure rate	Repair rate	e	v
1	0.0005	0.0027	0.8413	84.6505
2	0.0020	0.1060	0.9815	0.3184
3	0.0005	0.0265	0.9815	1.3282

The shape of the V surface as a function of buffer capacities N_1 and N_2 can be seen in Figure 7.1. The V surface decreases with N_1 and N_2 until a certain point then it starts to increase. Similar cases have been presented in [11, 82] and in Chapter 4 for 2 machine lines.

The CG algorithm was tested on this case to check if we can achieve the minimum V without any constraints. Figure 7.2 shows the iterations (represented by arrows) of the method. It can be seen that the step size α_k changes from iteration to iteration. Results are reported in Table 7.2.

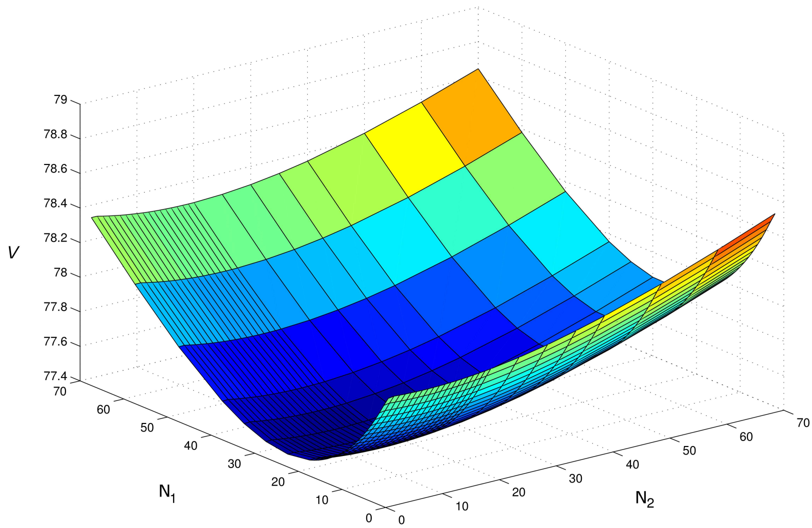


Figure 7.1: The surface of V for the test three machine line as a function of buffer capacities N_1 and N_2

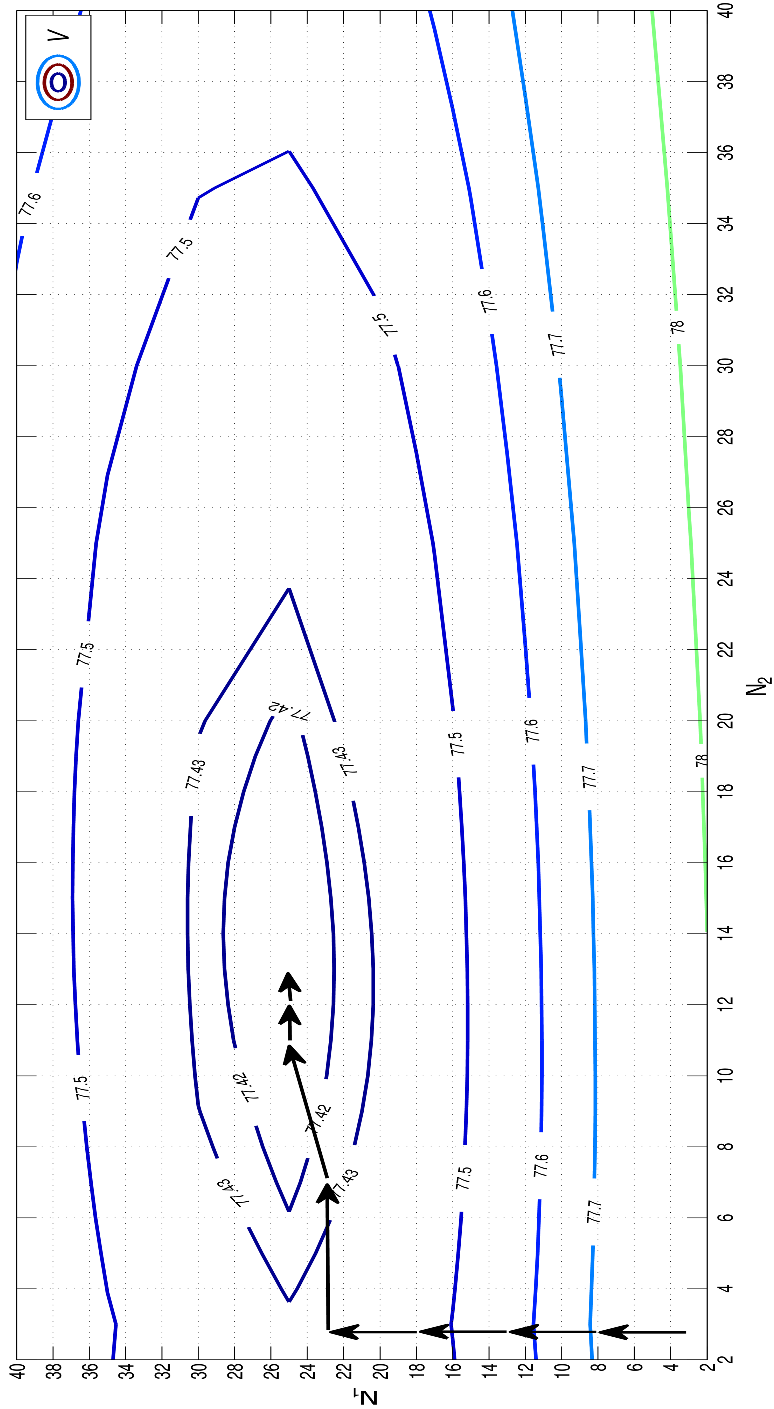


Figure 7.2: The application of the CG method, with arrows representing iterations

Table 7.2: Tabulated results regarding the application of the CG optimization on the tested case -unconstrained optimization

Iteration	Step Size	N_1	N_2	E	V
0		3	3	0.8159	77.9484
1	5	8	3	0.8171	77.7160
2	5	13	3	0.8183	77.5621
3	5	18	3	0.8193	77.4725
4	5	23	3	0.8202	77.4353
5	5	23	7	0.8209	77.4205
6	5	25	10	0.8217	77.4109
7	3	25	12	0.8221	77.4088
8	1	25	13	0.8222	77.4084

Finally, to check whether the BAP solution found by the CG algorithm that uses AGG method is the optimal solution, we can use the exact method proposed by Tan [84]. The neighbourhood points around the optimal solution were evaluated and reported in Table 7.3.

It can be seen that some of the neighbourhood points, evaluated by the exact method of Tan, have a lower V. This indicates that the V surfaces found by changing the buffers, in Tan's method and the AGG method, are not the same. The CG algorithm found the optimum buffer allocation from the V surface of the AGG method, however, this point does not correspond to an optimum buffer allocation on the V surface using Tan's method. Enhancing the accuracy of the approximate analytical methods for calculating V would improve the solution found by the CG algorithm, by finding V surfaces that are accurate and have the same behavior of the V surfaces found by the exact methods.

Table 7.3: Evaluation of the asymptotic variance rate for the neighbourhood points using Tan's method, around the optimal solution found with the approximate CG algorithm with AGG method

		N_2				
		11	12	13	14	15
	22	75.380	75.423	75.4649	75.5046	75.5428
	23	75.357	75.400	75.4425	75.4825	75.5211
	24	75.336	75.379	75.4215	75.4620	75.5009
N_1	25	75.315	75.359	**75.4020**	75.4428	75.4821
	26	75.296	75.341	75.3837	75.4249	75.4646
	27	75.278	75.323	75.3668	75.4084	75.4485
	28	75.262	75.307	75.3511	75.3931	75.4336

7.1.6 Constrained Optimization

A constrained optimization approach is applied to the three machine line two buffer cases reported in Section 6.3. The goal of the optimization is to solve the Dual problem i.e defined in 7.8. Figures [7.3 and 7.4] show contour graphs of V as a function of buffer capacities N_1 and N_2. The arrows show the iterations found by the optimization algorithm. These tests show that this adopted procedure moves in the correct direction of improvement. Further details about the performance measures at each step are reported in Tables [7.4 and 7.5].

Case 1: Constrained Primal BAP

Considering the Case 1 presented in Section 6.3 in which we try to solve the Primal BAP that states:

$$
\begin{aligned}
\text{minimize} \qquad & N^{\text{TOTAL}} = \sum_{i=1}^{K-1} N_i \\
\text{subject to} \qquad & E(N_1, \ldots, N_{K-1}) \geq 0.79 \\
& V(N_1, \ldots, N_{K-1} \leq 4.00 \\
& N_i \geq = 3, \text{i} = 1, \ldots, \text{K-1}
\end{aligned}
$$

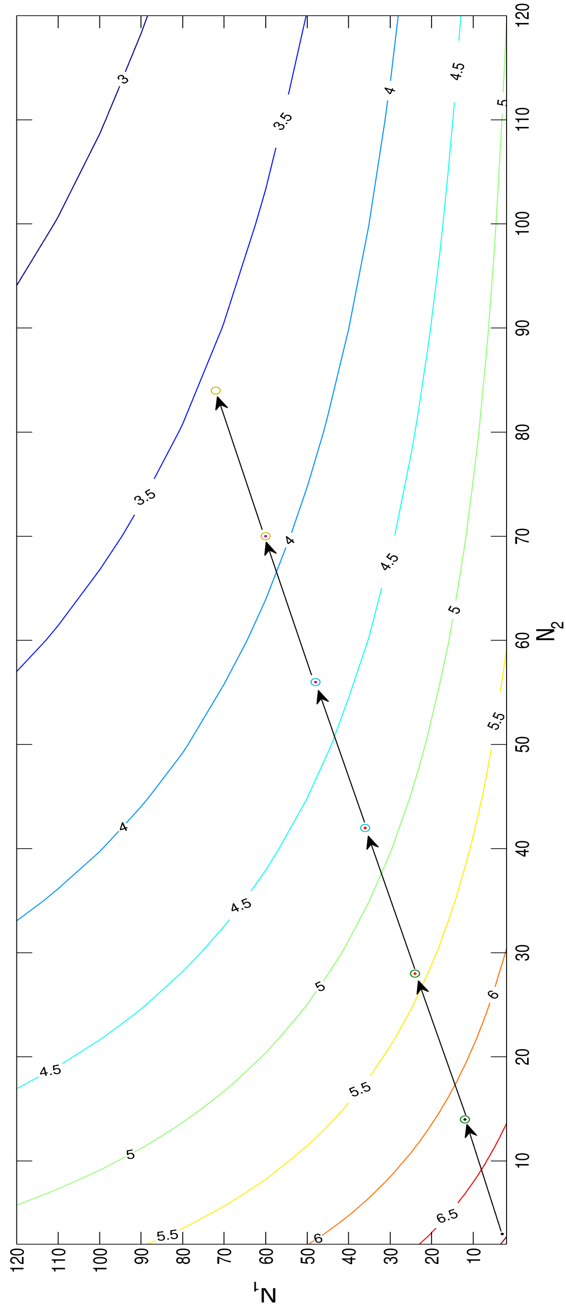


Figure 7.3: Optimization of Case 1 of the three machine line

Case 2: Constrained Dual BAP

Let's consider Case 2 presented in Section 6.3. This case can be solved using the Dual BAP. The formulation of the constrained problem is:

$$\begin{aligned} \text{minimize} \quad & V(N_1, \ldots, N_{K-1}) \\ \text{subject to} \quad & N^{\text{TOTAL}} \leq 111 \\ & N_i \geq 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned}$$

Figure 7.4 shows the result of applying the CG algorithm. The figure shows that only increasing N_2 reduces the output variability.

Table 7.4: Tabulated results regarding the application of the CG optimization on Case 1

Iteration	N_1	N_2	E	V
0	3	3	0.6969	6.9538
1	12	14	0.7247	6.1872
2	24	28	0.7487	5.424
3	36	42	0.7651	4.8188
4	48	56	0.7770	4.3215
5	60	70	0.7859	3.9051
6	72	84	0.7928	3.5531

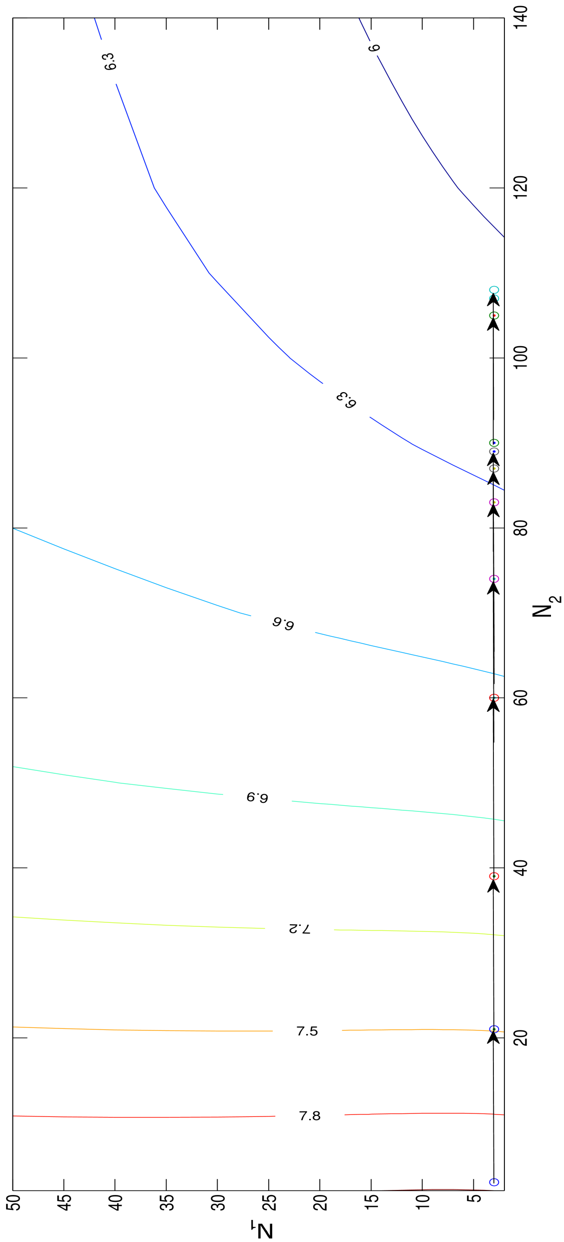


Figure 7.4: Optimization of Case 2 of the three machine line

To compare the effect of not considering the asymptotic variance rate on the BAP, we reformulate the problem as following:

$$\begin{aligned} \text{maximize} \quad & E(N_1, \ldots, N_{K-1}) \\ \text{subject to} \quad & N^{\text{TOTAL}} \leq 111 \\ & N_i \geq 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned}$$

Since we are not optimizing V, this problem can be solved using the gradient approach proposed by Gershwin and Schor [37]. Table 7.6 shows the result of the gradient approach without considering the output asymptotic variance rate V.

Table 7.6 shows that buffers are allocated differently when not considering the output variability in the problem. Although the average production rate E is higher (0.7852), the asymptotic variance rate is higher too (6.6822). The difference between the two BAP is about 9.3%. This means that overlooking the output variability in the BAP will generate suboptimal allocations in the system.

7.1.7 Service Level Optimization

Since the proposed gradient method was developed for non linear functions, it should work on the Service Level (SL) as well. Given a demand X after a certain period of T cycle times, the SL can be calculated by approximating the distribution of the output with a normal distribution [79]. Furthermore, since the proposed method minimizes the desired output, we need to consider $-SL$ as the function to minimize. The dual BAP formulation for the case under analysis can be expressed as:

$$\begin{aligned} \text{minimize} \quad & -SL(N_1, \ldots, N_{K-1}) \\ \text{subject to} \quad & N^{\text{TOTAL}} \leq 190 \\ & V(N_1, \ldots, N_{K-1}) \leq 5.2 \\ & E(N_1, \ldots, N_{K-1}) \geq 0.66 \\ & N_i \geq 3, \text{i} = 1, \ldots, \text{K-1} \end{aligned}$$

Table 7.5: Tabulated results regarding the application of the CG optimization on Case 2 of Section 6.3

Iteration	N_1	N_2	E	V
0	3	3	0.6973	8.0669
1	3	21	0.7277	7.4937
2	3	39	0.746	7.0398
3	3	60	0.7598	6.6429
4	3	74	0.7665	6.4364
5	3	83	0.7701	6.3223
6	3	87	0.7715	6.2755
7	3	89	0.7722	6.2529
8	3	90	0.7725	6.2418
9	3	105	0.7771	6.0902
10	3	107	0.7776	6.0718
11	3	108	0.7779	6.0628

Table 7.6: Tabulated results regarding the application of the traditional gradient method for optimization of Case 2 of Section 6.3 considering the average production rate

Iteration	N_1	N_2	E	V
0	3	3	0.6973	8.0669
1	16	24	0.7436	7.4140
2	29	43	0.7688	6.9601
3	33	51	0.7735	6.8843
4	38	60	0.7804	6.7552
5	42	67	0.7846	6.6870
6	43	68	0.7852	6.6822

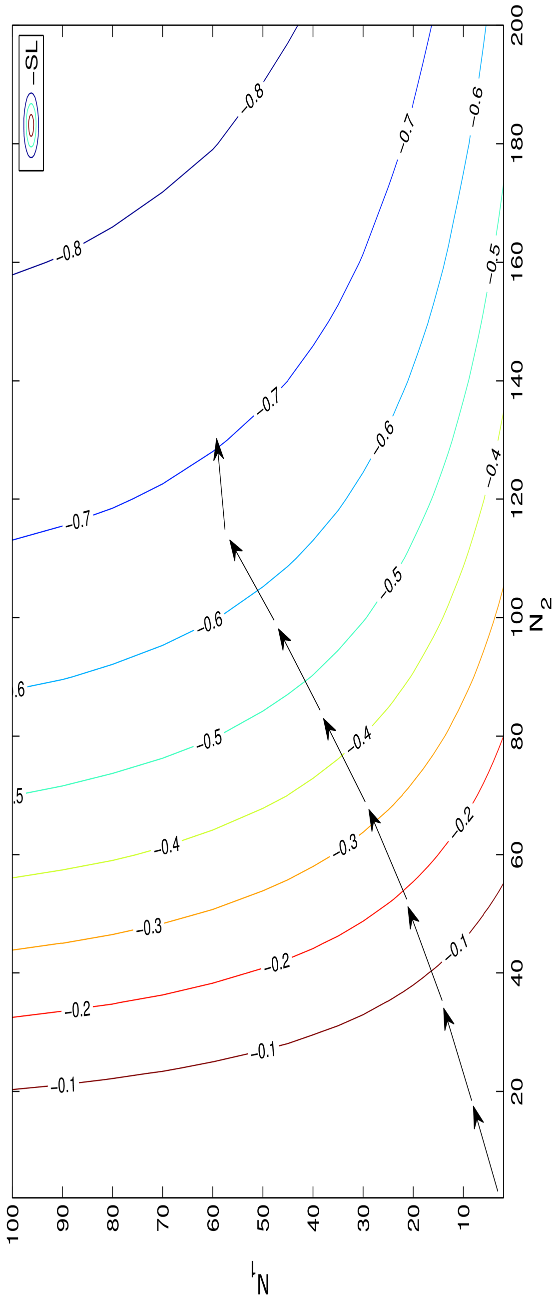


Figure 7.5: Optimization of Case 3 considering the SL

7.2 Conclusions

The mathematical modelling of the BAP that considers the output variability was proposed. Furthermore, this chapter proposes a new gradient optimization approach that can deal with V to optimize the new BAP. The new approach is based upon the Conjugate Gradient methods (CG) and it uses the Fletcher-Reeves-Polak-Ribière (FR-PR) approach to find new improvement direction. Results show that optimizing the BAP without considering the output variability could generate allocations that will result in higher variability in the system.

Results of applying the CG algorithm show that it can solve different constrained and unconstrained BAP. Although we have studied many other systems with the proposed optimization approach, we only reported a subset of cases that are representative of how practitioner could make use of the results proposed in this thesis. Finally the BAP was also optimized considering the service level SL by applying the same approach.

Table 7.7: Tabulated results regarding the application of the CG optimization on Case 3 of Section 6.3 considering the SL when the demand is 7900 products due time after 10,000 cycle times

Iteration	N_1	N_2	E	V	SL
0	3	3	0.6971	7.6996	0.0004
1	8	18	0.7288	7.1190	0.0100
2	14	35	0.7518	6.5598	0.0680
3	21	52	0.7678	6.1031	0.1839
4	29	68	0.7790	5.7489	0.3231
5	38	84	0.7876	5.4639	0.4591
6	47	99	0.7938	5.2535	0.5658
7	57	114	0.7987	5.0885	0.6501
8	59	131	0.8021	4.9620	0.7065

Chapter 8

Case Study: Levissima water bottling production line

The developed methods have been used to optimize the output of a real production system. The addressed problems are that of optimizing the buffer in a production line to reduce the output variability. In this Chapter, the main results achieved by means of the developed approaches are summarized, with a particular emphasis on the preliminary analysis which is needed while applying the developed methods to the real case. This Chapter reports an industrial problem that was analyzed in a Master's thesis work which has been carried out within the research group. More information and details can be found in Marconi [57].

The case presented in this chapter is a case from the beverage industry. The company has a bottling line for mineral water with different bottle sizes. The results presented in this thesis represent the period during which the production was concentrated on 0.5-litre bottles.

8.1 Company Profile

Levissima water is a brand produced by Gruppo Sanpellegrino S.p.A (http://www.sanpellegrino-corporate.it/) belonging to the multinational Nestlé organisation, which is the biggest and most important organization in Italy in the mineral waters and non-alcoholic beverage sector, with a large number of mineral waters, non-alcoholic aperitifs, soft drinks and cold teas. The group had a consolidated turnover of 877 millions of Euros in 2009, corresponding to a sales volume in the region of 3 billion bottles. Sanpellegrino group employs a workforce of 1.700.

The group products, S. Pellegrino mineral water in particular, are sold in more than 120 countries, through branches and distributors in all five continents. All 9 production units are located close to the mineral water springs in northern, central and south Italy, mainly in the alpine zone. The Group has 7 mineral waters: S.Pellegrino, Levissima, Acqua Panna, Nestlé Vera, Recoaro, S.Bernardo, Pejo [2].

Moreover, the Sanpellegrino Group has a complete range of products in terms of positioning and price band, which enables them to offer the consumer the product (water and soft drinks) and packaging (type and format) most suitable for their requirements. The brand positioning and market shares are growing year after year, as a result of the huge production investments and marketing and communications activities, among other factors. Their strongest brand Levissima accounts for more than 10% of the retail volume sales in Italy [1].

Other soft drinks include the famous Orangeade, Bitter Orange, Lemonade, Grapefruit, Lime, Chino Sanpellegrino, Belte, cold tea and the aperitifs Sanbitter and Gingerino.

Sanpellegrino has been working for years with the goal of safeguarding water resources through projects and initiatives that aim at enhancing the worth of this vital and precious asset, reducing its environmental impact and promoting a true water culture. Guaranteeing water a quality future is indeed the mission that Sanpellegrino pursues with commitment and consistency, firstly through the safeguarding of the areas in which it works [2].

8.2 Description of the Water Bottling Process

A short description of the considered water-bottling line will be presented first before going through with the output variability analysis. The line under consideration is dedicated for the production 0.5 liter bottles, with a nominal production rate of 38,000 bottles/hour.

The production line consists of seven stations that transform preforms into palletized burdens. The first station is a blow moulding machine, which is located in region A in Figure 8.1. The preforms are mounted on a special clips and heated to $127° - 128°$ by heating lamps along the way from the input location of the line until the arrival to the blow molding machine, the machine transforms the preheated preforms into empty bottles using pressure up to 25 bars. The blow molding machine has a production rate of 38,000 bottles per hour and it is the bottleneck machine of the line.

The empty bottles that are still attached to the clips go to the filling machine that fills them with the mineral water, Region B. The filled bottles are then controlled for defects in the bottles or incomplete filling before being sealed. Defective bottles are discarded automatically by an extractor.

Next, the bottles go to a labelling machine, Region C, which has a production rate of 41,000 bottles/hour. Labelled bottles are then controlled again for defects in the label, and at this point information regarding the production date, lot and exist time instance are added to the bottles.

Finished bottles later are grouped by a shrink bundler machine, Region D. The machine groups each 6 bottles together to create a small burden. The process continues to another shrink bundler machine, Region E, which groups every 4 small burdens into a larger burden. Thus each burden consists of 24 bottles.

The final stage of the production line consists of a palletizer machine, Figure 8.1 region F, that mounts 7 burdens over a pallet then sends it to a pallet shrink covering machine, Region H. This machine covers the complete pallet with a plastic cover. This is important for two reasons: it protects the products from contamination, and it ensures the stability of the pallet during transportation. The final pallets are stored in a big storage area that, approximately, can store 1000 pallets.

8.3 Problem Description

The main problem experienced by the company was the impact of the variability of the production output on the downstream product distribution operations. Due to the disturbances in the production, backlog frequency occurred resulting in delays in the product delivery to the various distribution centers and overall poor service level. The requirement of the company is to handle this problem by reducing the output variability of the line by exploiting the modularity of the conveyors as reconfigurability enablers, i.e. with minor capital investment. Before tackling the problem, technical data of the line is gathered and analyzed.

8.4 Technical Data of the Production Line

Levissima production line can be modeled as a production line composed of six machines and five buffers instead of seven machines and six buffers. The reason behind this decision is because the filling machine is much faster than the blow-molding machine, and the space between them is quite large. The two machines hence can be integrated into one machine with failure and repair times of both machines. In this section, the parameters of the line that are of major interest are reported.

8.4.1 Dimension of Buffers

The dimension of the buffers used between the machines are estimated based on the amount of products that would cause the machine before the buffer to be blocked. Table 8.1 shows the dimension of each of the buffers used. It should be noticed that the space between the blow molding machine and filling machine is very large and can be assumed to have infinite capacity.

8.4.2 Machine Reliability Parameters

The company uses a Computer Maintenance Management System (CMMS) for the collection of information about failures happening to the machines. The CMMS system also store such data in a database. From such database, it is possible to study the distributions of different machine failure times and the distribution of machine repair times. Analyzing in depth the failure and repair distributions for each machine

Table 8.1: The location and dimension of the buffers in Levissima bottling production line.

Location	Dimension (bottles)
Filling -Labeling	2496
Labeling - Shrink bundler 1	2004
Shrink Bundler 1 - Shrink Bundler 2	1500
Shrink Bundler 2 - Palletizer	3492
Palletizer - Shrink Bundler 3	4536

using the Kolmogorov - Smirnov test (with $\alpha = 0.12$), it is concluded that the up times and down times are geometrically distributed with $MTTF$ and $MTTR$ for each machine reported in table 8.2.

Table 8.2: MTTF and MTTR for the machines of the line

Machine	MTTF(hour)	MTTR(min)
Filling	7.96	13.8
Labeling	6.6	10.8
Shrink Bundler 1	12.11	6
Shrink Bundler 2	7.27	7.8
Palletizer	38.42	7.28
Shrink Bundler 3	81	22.8

8.4.3 The Production Rate of the Line

The production rate of the line is affected by the production rate of the bottleneck machine. In this case, the production rate corresponds to the first station, and it is equal to 38,000 bottles/ hour.

8.4.4 Preliminary Analysis of the Production Line

The analysis is concentrated on finding the buffer capacities that will optimize the total performance of the line in terms of average production rate E and the output asymptotic variance rate V. The optimization problem can be formulated without constraints, with limited constraints on the buffer capacities, and a set constraints on the buffer capacities and minimum E.

To apply the approximate analytical models for the evaluation of multi-stage production lines, the geometric distribution parameters that represent the probability of failure p and probability of repair r are calculated by:

$$p = \frac{ct}{MTTF} \tag{8.1}$$

$$r = \frac{ct}{MTTR} \tag{8.2}$$

where ct is the cycle time for each machine of the line, and it is equal to 0.0947 seconds. Since the cycle time is too small, it is more reasonable to apply a scaling of 50 bottles as 1 burden. By doing so the new cycle time becomes 4.736 seconds. Thus, the failure and repair rate parameters should be then calculated using equations (8.1) and (8.2) and the new cycle time of the machines. Table 8.3 shows the resulting p and r that will be used, while table 8.4 shows the buffer capacities after applying the scaling. Furthermore, the efficiency in isolation e and the asymptotic variance rate for isolated machines v are reported in table 8.5.

The output performance measures for Levissima's production line with current buffer configuration and using the aggregation method (AGG) is 0.9280 for the average production rate E and 17.4354 for the line's asymptotic variance rate V. The AGG method will be used for the analysis done for this case.

Table 8.5 shows that the Filling machine is the bottle neck machine and it has the highest v in isolation. Results from Section 6.3 Case 1 indicate that the system is mostly affected by the v of the bottle neck machine, so it is expected that the system will produce an output variability around 9.2 if supplied with infinite buffer. Furthermore, the system output asymptotic variance rate V would decrease with the increase of buffer capacities. Eventual possible options to reduce the output variability can be in one of the three possible forms:

Table 8.3: The probabilities of failure and repair for each machine of the line

Machine	p	r
Filling	1.653×10^{-4}	5.721×10^{-3}
Labeling	1.994×10^{-4}	7.310×10^{-2}
Shrink Bundler 1	1.087×10^{-4}	1.316×10^{-2}
Shrink Bundler 2	1.810×10^{-4}	1.084×10^{-2}
Palletizer	3.425×10^{-5}	2.6027×10^{-3}
Shrink Bundler 3	1.605×10^{-5}	3.463×10^{-3}

1. Optimize the buffer capacity, trying to minimize V while maintaining high E.

2. Reduce the MTTR of the system because the machine is in region A of Figure 4.6. Reducing the MTTR of the machine will have a double positive effect causing both e to increase and v to decrease. This can be achieved by additional training for the repair crew, or dedicating a repair crew for this machine as it has a critical effect on the system performance.

3. Again since the machine is in region A of Figure 4.6, procedures like machine improvement approaches for the Filling machine should drive the MTTF to be longer and that will cause e to increase and v to decrease.

8.5 Optimization of the Performance Measures of the Line

To minimize the output asymptotic variance rate of Levissima water bottling line, the CG gradient optimization method proposed in Chapter 7 was used in two different problem settings: unconstrained optimization and constrained optimization.

Table 8.4: The location and scaled dimension of the buffers in Levissima bottling production line.

Location	Dimension (50 bottle burdens)
Filling -Labeling	50
Labeling - Shrink bundler 1	41
Shrink Bundler 1 - Shrink Bundler 2	30
Shrink Bundler 2 - Palletizer	70
Palletizer - Shrink Bundler 3	91

8.5.1 Unconstrained Minimization of the Asymptotic Variance Rate of the Line

The unconstrained optimization of the dual problem tries to find the optimal buffer allocation that will generate the minimum output asymptotic variance rate V. The problem can be formally presented as:

$$\min \qquad V(N_1, \ldots, N_{K-1})$$

Table 8.6 presents the iterations performed using the conjugate gradient method. It can be seen even if the buffer capacities are increased by about 50%, the output variability is only reduced by about 9%. The system behaves the same way the three-machine line presented in Section 6.3 Case 1 i.e the larger the buffer capacities, the lower the asymptotic variance rate until arriving the asymptotic variance rate of the bottle neck machine. The decision in such situation is to consider improving the bottle neck machine reliability.

Table 8.5: The single machine performance measures e and v

Machine	e	v
Filling	0.9719	9.2462
Labeling	0.9735	6.8584
Shrink Bundler 1	0.9918	1.2165
Shrink Bundler 2	0.9824	3.3343
Palletizer	0.9969	0.5743
Shrink Bundler 3	0.9953	2.6352

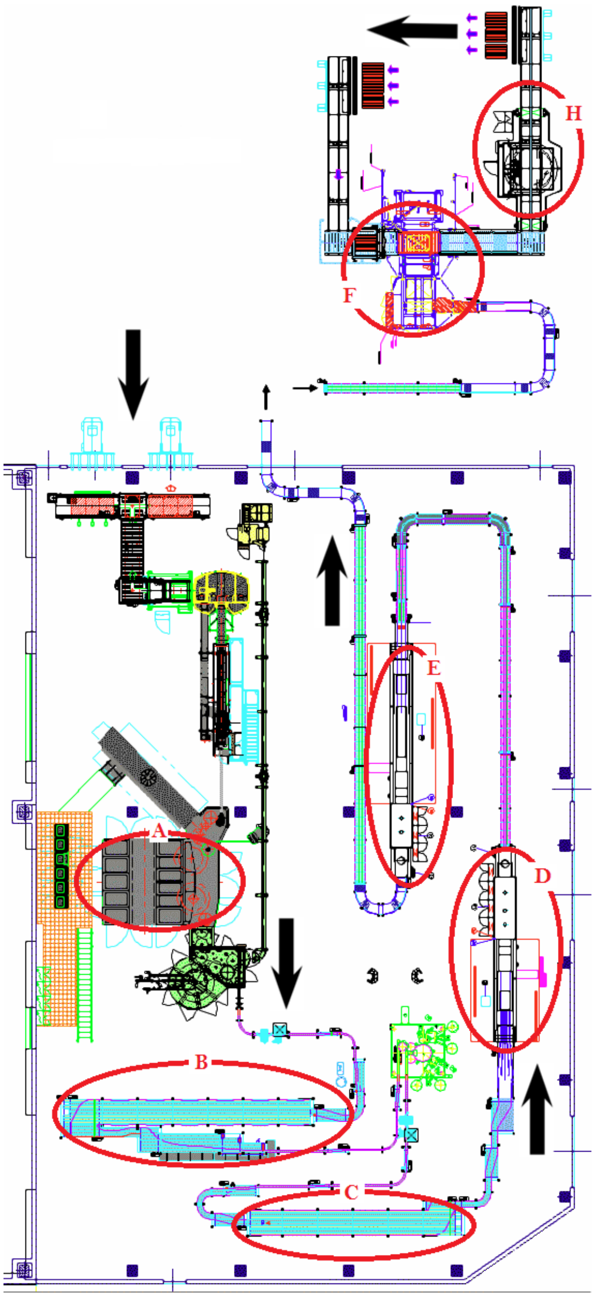


Figure 8.1: The layout of the production line of Levissima - black arrows represent the flow of parts

Table 8.6: The results of line optimization applying the gradient conjugate method- unconstrained optimization case

Iteration Number	Buffer sizes					% increase in Buffer sizes	E	V	Average Work in Process					
0	50	41	30	70	91		0.9279	17.4354	26.711	9.161	9.313	4.391	3.905
1	59	49	40	77	99	0.14	0.9298	17.0371	31.459	10.956	12.313	4.788	4.277
2	67	53	47	80	103	0.24	0.9311	16.7662	35.663	11.774	14.404	4.984	4.483
3	77	56	52	82	106	0.32	0.9321	16.5123	40.897	12.419	15.935	5.129	4.646
4	89	60	60	84	110	0.42	0.9335	16.1886	47.151	13.261	18.34	5.298	4.87
5	96	62	66	85	112	0.49	0.9343	15.993	50.786	13.689	20.108	5.395	4.994

8.5.2 Constrained Minimization of the Asymptotic Variance Rate of the Line

Since the problem of minimization of the asymptotic variance rate is unbounded, a useful modification is to add constraints on the buffer size to a certain limit. The problem can be formulated as following:

$$\begin{aligned} \text{minimize} \quad & V(N_1, \ldots, N_5) \\ \text{subject to} \quad & N^{\text{TOTAL}} = 282 \\ & N_i \geq N^{\text{MIN}} = 3, \text{i} = 1, \ldots, 5 \end{aligned}$$

This problem can be solved by the non-linear conjugate gradient method as well. To do so, we can start from the minimum buffer capacities, then add buffers to achieve the total buffer limit.

Table 8.7 shows the application of the conjugate gradient method. It shows that a reduction of about 5% in the asymptotic variance rate, and an increase of about 0.5% in the average throughput can be achieved without adding additional buffers to the production line. Finally, a comparison between the two adopted approaches is reported in Table 8.8.

Table 8.7: The results of line optimization applying the gradient conjugate method- constrained optimization case

Iteration Number	Buffer sizes					% increase in Buffer sizes	E	V	Average Work in Process				
0	3	3	3	3	3	15	0.9177	19.6395	1.5419	1.1812	1.243	0.9797	0.953
1	10	6	9	4	5	34	0.9191	19.3684	5.3357	1.8606	2.9643	1.0354	1.01
2	14	12	13	11	11	61	0.9205	19.1193	7.4958	3.193	4.129	1.4111	1.1826
3	22	17	21	15	16	91	0.9223	18.7648	11.7998	4.2981	6.4485	1.638	1.3373
4	37	25	35	20	23	140	0.9252	18.1753	19.8158	6.0543	10.5214	1.9411	1.5722
5	60	36	55	26	33	210	0.9289	17.371	31.9854	8.4407	16.3543	2.332	1.9379
6	82	47	74	32	43	278	0.9322	16.6085	44.5494	10.7804	22.447	2.7422	2.3387

Table 8.8: A summary of different optimization approaches

Prob.	E			V			Buffer levels		
	Curr.	Rec.	Δ%	Curr.	Rec.	Δ%	Curr.	Rec.	Δ%
Uncons	0.9279	0.9343	0.69%	17.4354	15.9930	-9.02%	282	421	49.29%
Cons		0.9322	0.46%		16.6085	-4.98%		282	0.00%

8.5.3 Conclusions

In this section we analysed the case study of Levissima, explaining in details how collected data from the production line can be used to evaluate the output variability of the system. Further investigation shows that applying the non-linear conjugate gradient method would be able to reduce the output variability by about -5% with no additional buffers and by about -9% with a 49% increase to the buffer capacities of the machines with a slight improvement to the buffer capacities.

Chapter 9

Managerial Insights, Conclusions and Recommendations

This work has presented a method for estimating the output variability of manufacturing systems. This chapter, will present the final part of the work in which guidelines and insights to production managers and practitioners will be presented. Then a complete conclusion that covers all parts of the work will be given. Finally, this chapter presents recommendations and suggestions for the future extensions of the work.

9.1 Managerial Insights

The managerial aspects of this work are various. The first aspect will discuss the effect of output variability on the service level of the manufacturing system. Then, insights on the buffer size effect on the variability will be discussed in depth. Manufacturing systems design considering the output variability will be discussed. Finally, Insights on machine reliability improvement vs. repair time reduction plans and their relationship to output variability will also be discussed.

Effect of Output Variability on the Service Level: Figures 1.4 and 1.8, show how much the asymptotic variance rate impacts the service level of the system. Other Figures such as 4.7 show that when the average demand is lower than average production rate at a certain time t, the system that exhibits higher correlation thus higher output variability negatively affect the service level. Although this result is intuitive, prior study of this phenomena at a manufacturing systems was not clear. However, the most important issue to consider is the fact that ignoring the output variability means losing control over the service level, causing extra costs in many different ways.

Buffer Size Effect on the Output Variability: Previous studies of Carrascosa and Tan have shown that the asymptotic variance rate can have a non-monotonic behaviour as a function of buffer capacities in two machine production systems. In this work, we found that this happens in longer lines as well, and arrived to the conclusion that the output variability of the system is affected mostly by the output variability of the bottle-neck machine. Thus, more attention to this machine should be given. From another point of view, increasing the buffer sizes for some systems could increase their output variability. This means that careful analysis of the system should be done before taking actions.

Manufacturing Systems Design: The analytical formulas derived in this work can help in manufacturing systems design phase. Tasks such as selecting machines, repair crew, and buffer sizes can be determined in order to optimally build the system to respect certain criteria. Managers should also notice that designing the system not considering output variability will result in suboptimal performance.

Machine Improvement Plans vs. Repair Time Reduction Actions: In chapter 3, Figure 4.6 shows that actions of increasing MTTF or decreasing MTTR do not always result in positive effects on the system's output variability. This means managers should be aware of this issue and should be very careful with their actions on how to improve machine reliability.

9.2 Conclusions

In the thesis, production models and quantitative methods to evaluate the output variability of manufacturing systems are provided. It has been shown that the problem addressed in this thesis is very relevant to both industrialists and researchers.

The first part of the thesis has been dedicated to the formalization of the considered problem and to the description of the modelled manufacturing systems. The second part presented a general analytical methodology for evaluating the output variability of single machines, two machine lines and small production systems modeled with a general Markovian structure.

The third part of the work tackled the approximate analytical evaluation of long manufacturing systems with multiple failure mode geometric machines. The proposed methods can estimate the output variability with an acceptable accuracy and speed. Insights on when to use these methods were given as well. The fourth part of the work studied the system behaviour of three machine systems and gave explanation on how the output variability is affected by the system parameters and machine characteristics. This helped proposing an optimization technique that can be applied for non-linear problems. The results show that this method succeeds in finding the optimal solution if the problem is bounded.

Finally, a case study was analyzed showing the process of evaluating the output variability starting from data collection process. Reconfigurations actions were suggested based on the application of the optimization procedure proposed.

9.3 Recommendations and Future Work

This work presented contributions in the evaluation of the output variability in manufacturing systems. The exact analytical evaluation was proposed for small manufacturing systems, while approximate methods were used to evaluate, and optimized longer production systems. The proposed approximate techniques depend on dividing the line into smaller components using decomposition or aggregation approaches. Extensions to the proposed approaches can be summarized as following:

- *The exact analytical evaluation of small manufacturing systems:* The proposed method can analyze manufacturing systems in discrete time that has a binary reward output vector. First, improvement would be extending the systems structure to general reward structure. Secondly, extending the proposed approach to machines with different processing times in continuous time. This could give more insights on the effect of the processing times on the output variability. The proposed approach for the evaluation of the asymptotic variance rate if from the family of state space approaches. These methods are not so computationally efficient when buffer capacities increase and machines are modeled with Phase-Type distribution or even multiple failure mode machines. An important extension would be deriving a new method that does not depend on the buffer size.

- *The approximate analytical evaluation of long manufacturing systems:* The different proposed approximate approaches give average error in the range of 5-16% (excluding CMT). This error can be furthermore reduced by working on the following issues:

  1. Regarding the AC method: The accuracy of the method should improve when using decomposition equations that are able to consider machine models very different from the geometric distribution.
  2. Regarding the Aggregation method: This method's accuracy might improve if we match the output process considering more moments. For example, we can match the first three moments of an output process, as a two-failure mode machine or as a three phase type machine.

- *The optimization of the production line considering output variability*: The work presented is considered a first step for the optimization of production systems considering the output variability. Future works can extend this approach for more production systems layouts.

Bibliography

[1] Italian water market far from saturation point. Online, August 2004. URL http://www.foodanddrinkeurope.com/Consumer-Trends/Italian-water-market-far-from-saturation-point.

[2] Fondazione Altagamma. Sanpellegrino. Online, 2010. URL http://touch.altagamma.it/index.php?sez_pagina=soci2&IdSocio=44#.

[3] T. Altiok and S. Jr Stidham. The allocation of interstage buffer capacities in production lines. *IIE Transactions*, 1993.

[4] A Angius, A Horvath, and M Colledani. Moments of cumulated output and completion time of unreliable general markovian machines. In *Proceeding of the 18th World Congress of the International Federation of Automatic Control (IFAC)*, Milan, Italy, September 2011.

[5] B. Baynat, J. A. Buzacott, and Y. Dallery. Multiproduct kanban-like control systems. *International Journal of Production Research*, 40(16):4225–4255, 2002. doi: 10.1080/00207540210146198. URL http://www.tandfonline.com/doi/abs/10.1080/00207540210146198.

[6] Gabriel R. Bitran and Devanath Tirupati. Multiproduct queueing networks with deterministic routing: Decomposition approach and the notion of interference. *Management Science*, 34(1):75–100, 1988. URL http://www.jstor.org/stable/2632187.

[7] A. M. Bonvik, C.E. Couch, and S.B. Gershwin. A comparison of production-line control mechanisms. *International Journal*

of Production Research, 35(3):789–804, 1997. doi: 10.1080/002075497195713. URL http://www.tandfonline.com/doi/abs/10.1080/002075497195713.

[8] P. Bremaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer, USA, 1999.

[9] J. A. Buzacott. Automatic transfer lines with buffer stocks. *International Journal of Production Research*, 5(3):183–200, 1967. doi: 10.1080/00207546708929751. URL http://www.tandfonline.com/doi/abs/10.1080/00207546708929751.

[10] John A. Buzacott and J. George Shanthikumar. Prentice Hall, Englewood Cliffs, N.J. ISBN 0138475679.

[11] M. Carrascosa. Variance of the output in a deterministic two-machine line. Master's thesis, Massachusetts Institute of Technology, Cambridge MA 02139, 1995.

[12] Chin-Tai Chen and John Yuan. Transient throughput analysis for a series type system of machines in terms of alternating renewal processes. *European Journal of Operational Research*, 155(1):178 – 197, 2004. ISSN 0377-2217. doi: 10.1016/S0377-2217(02)00838-X. URL http://www.sciencedirect.com/science/article/pii/S037722170200838X.

[13] S.Y. Chiang, C.T. Kuo, J.T. Lim, and S.M. Meerkov. Improvability of assembly systems i: problem formulation and performance evaluation. *Mathematical Problems in Engineering*, 6:359–393, 2000.

[14] Philippe Ciprut, Max-Olivier Hongler, and Yves Salama. Fluctuations of the production output of transfer lines. *Journal of Intelligent Manufacturing*, 11:183–189, 2000. ISSN 0956-5515. URL http://dx.doi.org/10.1023/A:1008942917166. 10.1023/A:1008942917166.

[15] M. Colledani and T. Tolio. A decomposition method to support the configuration / reconfiguration of production systems. *Annals of the CIRP*, 54, 2005.

[16] M. Colledani, F. Gandola, A. Matta, and T. Tolio. Performance evaluation of linear and non-linear multi-product multi-stage lines with unreliable machines and finite homogeneous buffers. *IIE Transactions*, 40(6):612–626, 2008. doi: 10.1080/

07408170701745345. URL http://www.tandfonline.com/doi/abs/10.1080/07408170701745345.

[17] M. Colledani, A. Matta, and T. Tolio. Analysis of the production variability in manufacturing lines. *ASME Conference Proceedings*, 2008(48357):381–390, 2008. doi: 10.1115/ESDA2008-59408.

[18] M. Colledani, M. Ekvall, T. Lundholm, P. Moriggi, A. Polato, and T. Tolio. Analytical methods to support continuous improvements at scania. *International Journal of Production Research*, 48(7): 1913–1945, 2010. doi: 10.1080/00207540802538039. URL http://www.tandfonline.com/doi/abs/10.1080/00207540802538039.

[19] M. Colledani, A. Matta, and T. Tolio. Analysis of the production variability in multi-stage manufacturing systems. *CIRP Annals - Manufacturing Technology*, 59(1):449 – 452, 2010. ISSN 0007-8506. doi: DOI:10.1016/j.cirp.2010.03.142. URL http://www.sciencedirect.com/science/article/pii/S0007850610001435.

[20] Marcello Colledani. *Production Logistics and Quality Performance in Manufacturing Systems.* PhD thesis, Politecnico di Milano, 2007.

[21] Marcello Colledani and Tullio Tolio. Performance evaluation of production systems monitored by statistical process control and off-line inspections. *International Journal of Production Economics*, 120(2):348 – 367, 2009. ISSN 0925-5273. doi: 10.1016/j.ijpe.2007.07.011. URL http://www.sciencedirect.com/science/article/pii/S0925527309000115. ¡ce:title¿Special Issue on Introduction to Design and Analysis of Production Systems¡/ce:title¿.

[22] Y. Dallery, R. David, and X.L. Xie. Approximate analysis of transfer lines with unreliable machines and finite buffers. *Automatic Control, IEEE Transactions on*, 34(9):943–953, sep 1989. ISSN 0018-9286. doi: 10.1109/9.35807.

[23] Yves Dallery and Stanley B. Gershwin. Manufacturing flow line systems: a review of models and analytical results. *Queueing Systems*, 12:3–94, 1992. ISSN 0257-0130. URL http://dx.doi.org/10.1007/BF01158636. 10.1007/BF01158636.

[24] A. C. Diamantidis, C. T. Papadopoulos *, and M. I. Vidalis. Exact analysis of a discrete material three-station one-buffer merge system with unreliable machines. *International Journal*

of Production Research, 42(4):651–675, 2004. doi: 10.1080/00207540310001602838. URL http://www.tandfonline.com/doi/abs/10.1080/00207540310001602838.

[25] C. Dincer and B. Deler. On the distribution of throughput of transfer lines. *The Journal of the Operational Research Society*, 51 (10):1170–1178, 2000. ISSN 01605682. URL http://www.jstor.org/stable/253929.

[26] WJ Doll and MA Vonderembse. The evolution of manufacturing systems: Towards the post-industrial enterprise. *Omega*, 19 (5):401 – 411, 1991. ISSN 0305-0483. doi: 10.1016/0305-0483(91)90057-Z. URL http://www.sciencedirect.com/science/article/pii/030504839190057Z.

[27] Jie Gao, Yinliang Yao, Valerie Zhu, Linyan Sun, and Lin Lin. Service-oriented manufacturing: a new product pattern and manufacturing paradigm. *Journal of Intelligent Manufacturing*, 22: 435–446, 2011. ISSN 0956-5515. URL http://dx.doi.org/10.1007/s10845-009-0301-y. 10.1007/s10845-009-0301-y.

[28] S B. Gershwin. Variance of the output of a tandem production system. In R. D. Onvural and I. F. Akyildiz, editors, *Proceedings of the Second International Conference on Queuing Networks with Finite Capacity*, 1993.

[29] S B. Gershwin. *Manufacturing Systems Engineering*. PTR Prentice Hall, USA, 1994.

[30] S. B. Gershwin and I. C. Schick. Modeling and analysis of three-state transfer lines with unreliable machines and finite buffers. *Operations Research*, 31(2):354–380, 1983.

[31] S.B. Gershwin. Design and operation of manufacturing systems: the control-point policy. *IIE Transactions*, 32(10):891–906, 2000. doi: 10.1080/07408170008967448. URL http://www.tandfonline.com/doi/abs/10.1080/07408170008967448.

[32] S.B. Gershwin and Y. Goldis. Efficient algorithms for transfer line design. Technical report, Massachusetts Institute of Technology, MIT Laboratory for Manufacturing and Productivity, 1995.

[33] S.B. Gershwin and L.M. Werner. An approximate analytical method for evaluating the performance of closed-loop flow systems with unreliable machines and finite buffers-part ii: large loops. May 2003.

[34] Stanley B. Gershwin. An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking. *Operations Research*, 35(2):291–305, 1987. URL http://www.jstor.org/stable/170701.

[35] Stanley B. Gershwin and Oded Berman. Analysis of transfer lines consisting of two unreliable machines with random processing times and finite storage buffers. *A I I E Transactions*, 13(1):2–11, 1981. doi: 10.1080/05695558108974530. URL http://www.tandfonline.com/doi/abs/10.1080/05695558108974530.

[36] Stanley B. Gershwin and Mitchell H. Burman. A decomposition method for analyzing inhomogeneous assembly/disassembly systems. *Annals of Operations Research*, 93:91–115, 2000. ISSN 0254-5330. URL http://dx.doi.org/10.1023/A:1018940310682. 10.1023/A:1018940310682.

[37] Stanley B. Gershwin and James E. Schor. Efficient algorithms for buffer space allocation. *Annals of Operations Research*, 93:117–144, 2000. ISSN 0254-5330. URL http://dx.doi.org/10.1023/A:1018988226612. 10.1023/A:1018988226612.

[38] Manish K. Govil and Michael C. Fu. Queueing theory in manufacturing: A survey. *Journal of Manufacturing Systems*, 18(3):214 – 240, 1999. ISSN 0278-6125. doi: 10.1016/S0278-6125(99)80033-8. URL http://www.sciencedirect.com/science/article/pii/S0278612599800338.

[39] W. K. Grassman. Means and variances in markov reward systems. In C. D. Meyer and R. J. Plemmons, editors, *Linear Algebra, Markov Chains and Queueing Models*, volume 48 of *The IMA Volumes in Mathematics and Its Applications*, pages 193–204. Springer-Verlag, New York, 1993.

[40] Xin-Feng He, Su Wu, and Quan-Lin Li. Production variability of production lines. *International Journal of Production Economics*, 107(1):78–87, 2007. ISSN 0925-5273. doi: 10.1016/j.ijpe.2006.05.014. Special Section on Building Core-Competence through Operational Excellence.

[41] S. Helber. Performance analysis of flow lines with non-linear flow of material. lecture notes in economics and mathematical systems. 473, 1999.

[42] Kevin B. Hendricks. The output processes of serial production lines of exponential machines with finite buffers. *Operations Research*, 40(6):1139–1147, 1992. ISSN 0030364X.

[43] Kevin B. Hendricks and John O. McClain. The output processes of serial production lines of general machines with finite buffers. *Management Science*, 39(10):1194–1201, 1993. ISSN 00251909.

[44] A Horvath. Some results on discrete markovian reward models. Techical report, Turin University, July 2010.

[45] Robert R. Inman. Empirical evaluation of exponential and independence assumptions in queueing models of manufacturing systems. *Production and Operations Management*, 8(4):409–432, 1999. ISSN 1937-5956. doi: 10.1111/j.1937-5956.1999.tb00316.x. URL http://dx.doi.org/10.1111/j.1937-5956.1999.tb00316.x.

[46] Shin-Ichi Itoh. Customer-oriented manufacturing. *International Journal of Human Factors in Manufacturing*, 1(4):365–370, 1991. ISSN 1522-7111. doi: 10.1002/hfm.4530010407. URL http://dx.doi.org/10.1002/hfm.4530010407.

[47] M. A. Jafari and J. G. Shanthikumar. Exact and approximate solutions to two-stage transfer lines with general uptime and downtime distributions. *IIE Transactions*, 19(4):412–420, 1987. doi: 10.1080/07408178708975414. URL http://www.tandfonline.com/doi/abs/10.1080/07408178708975414.

[48] Muris Lage Junior and Moacir Godinho Filho. Variations of the kanban system: Literature review and classification. *International Journal of Production Economics*, 125(1):13 – 21, 2010. ISSN 0925-5273. doi: 10.1016/j.ijpe.2010.01.009. URL http://www.sciencedirect.com/science/article/pii/S0925527310000198.

[49] H. Le Bihan and Y. Dallery. Homogenization techniques for the analysis of production lines with unreliable machines having different speeds. *Eur. J. Contr.*, 3(3), 1997.

[50] Jingshan Li. Modelling and analysis of complex production systems. In *Proceedings of the fourth Aegean International Conference on Analysis of Manufacturing Systems*, pages 203–212, Samos Island, Greece, 2003. URL http://www.icsd.aegean.gr/aic2003/Papers/Li.pdf.

[51] Jingshan Li and S.M. Meerkov. Production variability in manufacturing systems: Bernoulli reliability case. *Annals of Operations Research*, 93(1):299–324, 2000. ISSN 1572-9338. doi: 10.1023/A:1018928007956.

[52] Jingshan Li and S.M. Meerkov. *Production Systems Engineering*. Springer, 2009.

[53] Na Li, Li Zheng, and Quan-Lin Li. Performance analysis of two-loop closed production systems. *Computers & Operations Research*, 36(1):119 – 134, 2009. ISSN 0305-0548. doi: 10.1016/j.cor.2007.07.008. URL `http://www.sciencedirect.com/science/article/pii/S0305054807001451`. ¡ce:title¿Part Special Issue: Operations Research Approaches for Disaster Recovery Planning¡/ce:title¿.

[54] J.-T. Lim, S.M. Meerkov, and F. Top. Homogeneous, asymptotically reliable serial production lines: theory and a case study. *Automatic Control, IEEE Transactions on*, 35(5):524 –534, may 1990. ISSN 0018-9286. doi: 10.1109/9.53518.

[55] Michael Manitz and Horst Tempelmeier. The variance of inter-departure times of the output of an assembly line with finite buffers, converging flow of material, and general service times. *OR Spectrum*, pages 1–19, 2010. ISSN 0171-6468. URL `http://dx.doi.org/10.1007/s00291-010-0216-1`. 10.1007/s00291-010-0216-1.

[56] John Mapes, Marek Szwejczewski, and Colin New. Process variability and its effect on plant performance. *International Journal of Operations and Production Management*, 20(7):792 –808, 2000.

[57] M. Marconi. Studio sulla varianza della produzione di linee a trasferta. Masters thesis, Politecnico di Milano, Faculta di Ingegneria Industriale, 2007.

[58] Andrea Matta and Quirico Semeraro, editors. *Design of Advanced Manufacturing Systems: Models for Capacity Planning in Advanced Manufacturing Systems*. Springer Netherlands, 2005. doi: 10.1007/1-4020-2931-4.

[59] Andrea Matta, Yves Dallery, and Maria Di Mascolo. Analysis of assembly systems controlled with kanbans. *European Journal of Operational Research*, 166(2):310 – 336, 2005. ISSN 0377-2217. doi:

10.1016/j.ejor.2003.09.035. URL http://www.sciencedirect.com/science/article/pii/S0377221704001250.

[60] S.M. Meerkov, N. Shimkin, and Liang Zhang. Transient behavior of two-machine geometric production lines. *Automatic Control, IEEE Transactions on*, 55(2):453 –458, feb. 2010. ISSN 0018-9286. doi: 10.1109/TAC.2009.2036328.

[61] G J. Miltenburg. Variance of the number of units produced on a transfer line with buffer inventories during a period of length t. *Naval Research Logistics*, 34(6):811–822, 1987. ISSN 1520-6750. doi: 10.1002/1520-6750(198712)34:6⟨811::AID-NAV3220340606⟩3.0.CO;2-Z.

[62] J. Miltenburg. U-shaped production lines: A review of theory and practice. *International Journal of Production Economics*, 70(3):201 – 214, 2001. ISSN 0925-5273. doi: 10.1016/S0925-5273(00)00064-5. URL http://www.sciencedirect.com/science/article/pii/S0925527300000645.

[63] Eginhard J. Muth. The reversibility property of production lines. *Management Science*, 25(2):pp. 152–158, 1979. ISSN 00251909. URL http://www.jstor.org/stable/2630617.

[64] Christrian Nalon. Analisi linee produzione con distribuzioni dei tempi di guasto e riparazione qualsiasi. Master's thesis, Politecnico di Milano, 2008.

[65] Y. Narahari and N. Viswanadham. Transient analysis of manufacturing systems performance. *Robotics and Automation, IEEE Transactions on*, 10(2):230 –244, apr 1994.

[66] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer, 2nd edition edition, 2006.

[67] Judy D. Olian and Sara L. Rynes. Making total quality work: Aligning organizational processes, performance measures, and stakeholders. *Human Resource Management*, 30(3):303–333, 1991. ISSN 1099-050X. doi: 10.1002/hrm.3930300303. URL http://dx.doi.org/10.1002/hrm.3930300303.

[68] Phillip F. Ostwald and Jairo Munoz. *Manufacturing Processes And System*. John Wiley and Sons, 1997. URL http://www.knovel.com/web/portal/browse/display?_EXT_KNOVEL_DISPLAY_bookid=2891&VerticalID=0.

[69] J. Ou and S B. Gershwin. The variance of the lead time of a two machine transfer line with a finite buffer. Techical report LMP-90-028, Laboratory for Manufacturing and Productivity, MIT, 1989.

[70] Chrissoleon T. Papadopoulos, Michael E. J. O'Kelly, Michael J. Vidalis, and Diomidis Spinellis. *Analysis and design of discrete part production lines.* Springer Optimization and Its Applications 31. New York, NY: Springer. xx, 279 p. EUR 80.20; SFR 124.50 , 2009. doi: 10.1007/978-0-387-89494-2.

[71] H.T. Papadopoulos and C. Heavey. Queueing theory in manufacturing systems analysis and design: A classification of models for production and transfer lines. *European Journal of Operational Research*, 92(1):1 – 27, 1996. ISSN 0377-2217. doi: 10.1016/0377-2217(95)00378-9. URL http://www.sciencedirect.com/science/article/pii/0377221795003789.

[72] Levantesi R. *Analysis of multiple loop assembly/disassembly networks.* Ph.d thesis, Politecnico di Milano, Milan, Italy, 2001.

[73] H. Robert and Chenhall. Reliance on manufacturing performance measures, total quality management and organizational performance. *Management Accounting Research*, 8(2):187 – 206, 1997. ISSN 1044-5005. doi: 10.1006/mare.1996.0038. URL http://www.sciencedirect.com/science/article/pii/S1044500596900384.

[74] IHSAN SABUNCUOGLU, ERDAL EREL, and A. GURHAN KOK. Analysis of assembly systems for interdeparture time variability and throughput. *IIE Transactions*, 34(1):23–40, 2002. doi: 10.1080/07408170208928847. URL http://www.tandfonline.com/doi/abs/10.1080/07408170208928847.

[75] Chuan Shi and Stanley B. Gershwin. An efficient buffer design algorithm for production line profit maximization. *International Journal of Production Economics*, 122(2):725 – 740, 2009. ISSN 0925-5273. doi: 10.1016/j.ijpe.2009.06.040. URL http://www.sciencedirect.com/science/article/pii/S0925527309002382.

[76] William J. Stewart. *Probability, Markov chains, queues and simulation: the mathematical basis of performance modeling.* Princeton University Press, 1st edition, 2009.

[77] B. Tan. Variance of the throughput of an n-station production line with no intermediate buffers and time dependent failures. *European Journal of Operational Research*, 101(3):560–576, 1997. ISSN 0377-2217. doi: 10.1016/S0377-2217(96)00191-9.

[78] B. Tan. Agile manufacturing and management of variability. *International Transactions in Operational Research*, 5:375–388, 1998. doi: 10.1016/S0969-6016(98)00024-0.

[79] B. Tan. Effects of variability on the due-time performance of a continuous materials flow production system in series. *International Journal of Production Economics*, 54(1):87–100, 1998. ISSN 0925-5273. doi: 10.1016/S0925-5273(97)00132-1.

[80] B. Tan. An analytical formula for variance of output from a series-parallel production system with no interstation buffers and time-dependent failures. *Mathematical and Computer Modelling*, 27 (6):95–112, 1998. ISSN 0895-7177. doi: 10.1016/S0895-7177(98) 00031-4.

[81] B. Tan. Variance of the output as a function of time: Production line dynamics. *European Journal of Operational Research*, 117 (3):470–484, 1999. ISSN 0377-2217. doi: 10.1016/S0377-2217(98) 00266-5.

[82] B. Tan. Asymptotic variance rate of the output in production lines with finite buffers. *Annals of Operations Research*, 93(1):385–403, 2000. ISSN 1572-9338. doi: 10.1023/A:1018992327521.

[83] B. Tan and S. Yeralan. Analysis of multistation production systems with limited buffer capacity part 2: The decomposition method. *Mathematical and Computer Modelling*, 1997.

[84] Baris Tan. State-space modeling and analysis of pull-controlled production systems. In S.B. Gershwin, Yves Dallery, C.T. Papadopoulos, and J. MacGregor Smith, editors, *Analysis and Modeling of Manufacturing Systems*, Operations Research Management Science. KAP, 2002.

[85] T. Tolio, A. Matta, and F. Jovane. A method for performance evaluation of automated flow lines. *CIRP Annals - Manufacturing Technology*, 47(1):373 – 376, 1998. ISSN 0007-8506. doi: DOI: 10.1016/S0007-8506(07)62854-6.

[86] T. Tolio, D. Ceglarek, H.A. ElMaraghy, A. Fischer, S.J. Hu, L. Laperriere, S.T. Newman, and J. Vancza. Species co-evolution of products, processes and production systems. *CIRP Annals - Manufacturing Technology*, 59(2):672 – 693, 2010. ISSN 0007-8506. doi: 10.1016/j.cirp.2010.05.008. URL http://www.sciencedirect.com/science/article/pii/S0007850610001952.

[87] Tullio Tolio, Andrea Matta, and S.B. Gershwin. Analysis of two-machine lines with multiple failure modes. *IIE Transactions*, 34:51–62, 2002. ISSN 0740-817X. URL http://dx.doi.org/10.1023/A:1019293215459. 10.1023/A:1019293215459.

[88] Tullio Tolio, Andrea Matta, and S.B. Gershwin. Analysis of two-machine lines with multiple failure modes. *IIE Transactions*, 34:51–62, 2002. ISSN 0740-817X. URL http://dx.doi.org/10.1023/A:1019293215459. 10.1023/A:1019293215459.

[89] F. Véricourt and S.B. Gershwin. Performance evaluation of a make-to-stock production line with a two-parameter-per-machine policy: the control point policy. *IIE Transactions*, 36(3):221–236, 2004. doi: 10.1080/07408170490274197. URL http://www.tandfonline.com/doi/abs/10.1080/07408170490274197.

[90] P. Welch. *The computer modeling handbook*, chapter The statistical analysis of simulation results, pages 268–328. Academic Press, New York, 1983.

[91] Sencer Yeralan and Eginhard J. Muth. A general model of a production line with intermediate buffer and station breakdown. *IIE Transactions*, 19(2):130–139, 1987. doi: 10.1080/07408178708975379. URL http://www.tandfonline.com/doi/abs/10.1080/07408178708975379.