The (α, β)-Precision Theory for Production System Monitoring and Improvement

by

Kang Liu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in The University of Michigan
2021

Doctoral Committee:

Professor Semyon M. Meerkov, Chair
Professor Roman Kapuscinski
Professor S. Sandeep Pradhan
Professor Lei Ying

Kang Liu

kangliu@umich.edu

ORCID iD: 0000-0003-2002-984X

© Kang Liu 2021

DEDICATION

This dissertation is dedicated to my parents Qingyong Liu and Xiuying Sun.

ACKNOWLEDGEMENTS

I would like to thank my Ph.D. advisor Prof. Semyon M. Meerkov, my committee members Prof. Roman Kapuscinski, Prof. S. Sandeep Pradhan, Prof. Lei Ying, my collaborators on the work related to this dissertation Dr. Pooya Alavian, Prof. Yongsoon Eun, Prof. Liang Zhang, and my colleagues in other publication Prof. Anouck Girard, Prof. Ilya Kolmanovsky and Dr. Nan Li. I would also like to thank my parents Qingyong Liu and Xiuying Sun.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

In the field of production system engineering, machine parameters, such as Mean Time Between Failures ($MTBF$), Mean Time To Repair ($MTTR$), machine quality parameter (q), and machine cycle time (τ), are widely used in quantitative methods for production system performance analysis, continuous improvement, and design. Unfortunately, the literature offers no methods for evaluating the smallest number of measurements necessary and sufficient to calculate reliable estimates of these parameters and the induced estimates of system performance metrics, such as machine efficiency (e), throughput (TP), quality parts throughput (TP_q), production lead time (LT), and work-in-process (WIP). This dissertation is intended to provide such a method. The approach is based on introducing the notation of (α, β)-precise estimates, where α characterizes the estimate's accuracy and β its probability.

Using this notion, the smallest number, $n_T^*(\alpha, \beta)$, of up- and downtime measurements necessary and sufficient to ensure (α, β)-precise estimates of $MTBF$ and $MTTR$ is calculated, and a probabilistic upper bound of the observation time required to collect $n_T^*(\alpha, \beta)$ measurements is derived.

The $MTBF$ and $MTTR$ are used to calculate production systems performance metrics e, TP, LT, WIP, which are necessary for managing production systems and for evaluating effectiveness of potential continuous improvement projects. This

dissertation evaluates the induced precision of these performance metrics estimates, based on (α, β)-precise estimates of $MTBF$ and $MTTR$. An inverse problem, i.e., calculating the smallest number of machines' up- and downtime measurements to ensure these performance metrics estimates with a desired precision, is also solved.

Along with $MTBF$ and $MTTR$, the machine quality parameter q, which represents the probability that a part produced is non-defective and $1-q$ the probability that it is defective, is used to evaluate quality parts throughput TP_q. This dissertation calculates the smallest number of parts quality measurements to ensure (α_q, β_q)-precise estimate of q, evaluates the induced precision of TP_q estimates, and presents solution to the inverse problem concerning q and TP_q, i.e., calculating the smallest number of parts quality measurements to ensure estimates of TP_q with a desired precision.

The (α, β)-Precision Theory is also compared with other probabilistic method, which can be used for evaluating the critical numbers. Specifically, we consider the Markov inequality, Chebyshev inequality and the simulation approach. It is shown in this work that these classical probability inequalities can only give estimates of the critical numbers that are much larger than their real values, which are too conservative and unrealistic to be implemented in practice. The simulation method also has several disadvantages compared with the theory, for example, large computation time complexity and the necessity to repeat these calculations if the parameters of the systems are changed.

In addition, this dissertation applies the (α, β)-Precision Theory to the study of production systems with cycle overrun. The cycle overrun takes place, for instance,

in automated machines with a constant part processing time, τ, and manual loading/unloading operations, which may have a random overrun in their duration. In this dissertation, the methods to obtain reliable estimates of cycle overrun, and the modeling, analysis, improvability and bottleneck identification of such systems are presented as well. Finally, the dissertation presents a case study motivated by an automotive transmission machining line.

CHAPTER 1

Introduction

1.1 Goal and Approach of the Dissertation

In the field of production system engineering, machine parameters, such as Mean Time Between Failures ($MTBF$), Mean Time To Repair ($MTTR$), machine quality parameter (q), and machine cycle time (τ), are widely used in quantitative methods for production system performance analysis, continuous improvement, and design (*Law and Kelton* (1991); *Viswanadham and Narahari* (1992); *Askin and Standridge* (1993); *Buzacott and Shanthikumar* (1993); *Papadopolous et al.* (1993); *Gershwin* (1994); *Perros* (1994); *Altiok* (1997); *Jerry* (2005); *Li and Meerkov* (2009); *Papadopoulos et al.* (2009); *Curry and Feldman* (2010); *Altiok and Melamed* (2010)). Unfortunately, the literature offers no methods for evaluating the smallest number of measurements necessary and sufficient to calculate reliable estimates of these parameters and the induced estimates of system performance metrics, such as machine efficiency (e), throughput (TP), quality parts throughput (TP_q), production lead time (LT), and work-in-process (WIP). The law of large numbers guarantees that a sufficiently large number of measurements leads to a sufficiently precise estimate. However, collecting a large number of measurements requires a long observation time. This is acceptable for off-line evaluation of machine parameters and system performance metrics, where historical data could be used. For on-line applications (e.g., those in the framework of Industry 4.0 and Smart Manufacturing, see *Kagermann et al.* (2013); *Liao et al.* (2017)), the observation period must be as short as possible in order to capture the data necessary for real-time utilization.

The goal of this dissertation is to provide methods for evaluating the smallest number of measurements for calculating reliable estimates of machine parameters and system performance metrics.

The approach for evaluating the smallest number of measurements is based on the notion of (α, β)-precise estimates, where α characterizes the estimate's accuracy and β its probability. Using this notion, we evaluate the smallest number of machine measurements to ensure (α, β)-precise estimates of machine parameters and induced system performance metrics.

In addition, the topic of production systems with unreliable machines and cycle overrun is studied in this dissertation. The cycle overrun may takes place, for instance, in automated operations with a constant part processing time, τ, which may have a random component in their duration. Another scenario is typical in assembly operations, where a fixed cycle time is imposed by operational conveyors, and the overrun is enabled by push-buttons, offering the operator a possibility to occasionally stop the conveyor in order to complete the job with the desired quality. Given that the current literature offers no analytical methods for analysis and improvement of production systems with cycle overrun and taking into account that these systems are encountered in practice, developing such methods is of importance. Specifically, in this dissertation, we intend to provide the modeling, analysis, improvability and bottleneck identification of production systems with cycle overrun, as well as apply the (α, β)-Precision Theory to the cycle overrun parameter estimation.

1.2 Main Results of the Dissertation

The the main results obtained in this dissertation are as follows:

- The smallest number of machine up- and downtime measurements, denoted as $n_T^*(\alpha, \beta)$, which is necessary and sufficient to ensure (α, β)-precise estimates of $MTBF$ and $MTTR$, is calculated, and a probabilistic upper bound of the observation time required to collect $n_T^*(\alpha, \beta)$ measurements is derived.

- Based on the (α, β)-precise estimates of $MTBF$ and $MTTR$, we evaluated the induced precision of performance metrics estimates $\widehat{e}$, $\widehat{TP}$, $\widehat{LT}$, and $\widehat{WIP}$, which is quantified by a pair of two numbers, (α_X, β_X), where X is in the set $\{e, TP, LT, WIP\}$. Specifically, the values of (α_X, β_X) for each performance metrics are calculated as functions of (α, β).

- In addition to the induced precision problem, an inverse problem, i.e., calculating the smallest numbers of machines' up- and downtime measurements to ensure estimates $\widehat{e}$, $\widehat{TP}$, $\widehat{LT}$, and $\widehat{WIP}$ with a desired precision (γ, δ), is solved.

- The machine quality parameter q, which represents the probability that a part produced is non-defective and $1-q$ the probability that it is defective, is used to evaluate quality parts throughput TP_q. The smallest number of parts quality measurements, denoted as $n_q^*(\alpha_q, \beta_q)$, to ensure (α_q, β_q)-precise estimate of q, is calculated, the induced precision of TP_q estimate, quantified by $(\alpha_{TP_q}, \beta_{TP_q})$, is evaluated, and solution to the inverse problem concerning TP_q, i.e., calculating the smallest number of parts quality measurements to ensure estimates of TP_q with a desired precision (γ, δ), is presented.

- Based on the solutions to the inverse problems, we provided qualitative analysis of the variability property for e, TP, LT, WIP and TP_q.

- Since the estimates of machine parameters are random variables, their statistical characterization can also be obtained using the classical probabilistic inequalities, namely, Markov and Chebyshev inequalities, as well as numerical simulations. We compared the smallest numbers of measurements evaluated using these tools with $n_T^*(\alpha, \beta)$ and $n_q^*(\alpha_q, \beta_q)$ calculated using the (α, β)-Precision Theory, and illustrated that the Markov and Chebyshev inequalities can only give approximations of the smallest numbers of measurements, which are significantly larger compared with those obtained using our method. In terms of the simulation method, we showed that it has several disadvantages, for instance, large computation time complexity, and the necessity to repeat these calculations if the parameters of the systems are changed.

- The methods for modeling, analysis, improvability, bottleneck identification, and parameter estimation of the production systems with cycle overrun are developed.

- A case study, which applies the methods developed to the throughput improvement of a production system motivated by an automotive transmission case machining line, is carried out.

1.3 Literature Review

Although no analytical results, addressing the smallest number of measurements required for calculating the estimates of machine parameters $MTBF$, $MTTR$, q, τ, and the estimtes of system performance metrics e, TP, TP_q, LT, WIP, with the desired precision, are available in the literature, similar problems, concerning other parameters and performance metrics, have been discussed in recent years. Specifically, *Muchiri et al.* (2014) used simulations to analyze the probabilistic behavior of manufacturing equipment under corrective and preventative maintenance activities. *Zhou et al.* (2014) proposed a data-driven framework utilizing case-based reasoning to achieve online product quality estimate in industrial plants. *Yu and Matta* (2016) proposed a statistical framework to increase the accuracy of performance metrics measurements, leading to improved bottleneck identification. *Hao et al.* (2017) modeled the interaction of tool wear and product quality degradation by a continuous-time stochastic system, and proposed a Bayesian framework, which incorporates real-time quality measurements, to estimate residual life of manufacturing systems and product quality. *Hwang et al.* (2017) proposed a production system performance measurement process applicable in the Internet of Things environment. *Kontar et al.* (2017) studied the estimation of key performance indicators of manufacturing systems, using a multi-output Gaussian process model. *Saez et al.* (2018)

designed a real-time system efficiency assessment framework using Internet of Things solutions, and applied it to a fully automated manufacturing system with robots and CNC machines. *Chehade and Shi* (2019) proposed sensor fusion method for statistical hypothesis testing to achieve online machine performance evaluation. *Chhetri et al.* (2019) proposed a digital twin solution, which incorporates side-channel sensor information, such as acoustic and magnetic signals, to localize manufacturing systems' anomalous faults and infer the product quality in real-time. *Khatab et al.* (2019) developed a method to determine an optimal inspection cycle for a deteriorating single-machine production system. *Gopalswamy and Uzsoy* (2019) proposed a data-driven refinement approach to improve production system performance estimates under model uncertainties. *Lin et al.* (2019) developed an approach to evaluate manufacturing systems performance metrics based on synergetic between analytical and simulation techniques. *Chen and Wang* (2019) presented a method for approximating the marginal probability distribution of work-in-process within multi-product-type, multi-stage, multi-parallel-machine manufacturing systems. *Fang et al.* (2020) proposed a novel deep neural network structure to estimate the jobs remaining time and achieved higher estimate accuracy than the existing machine learning models. *Müller et al.* (2020) proposed methods to quantify and measure the sequence stability in production system and evaluated different performance indicators. *Schneckenreither et al.* (2020) used a neural network to dynamically estimate the production lead time and thus determine the system release policy.

As far as the study of production system with cycle overrun is concerned, it should be pointed out that, although the literature offers no analytical methods for analysis and improvement of such systems, some of the related issues have been discussed in manufacturing and automation engineering literature. Specifically, *Morrison and Martin* (2007) developed practical methods for approximating random cycle time of manufacturing systems modeled by a G/G/M-queue. *Nadarajah and Kotz* (2008) provided the cycle time distribution formula to characterize the cycle underrun and overrun, where the cycle time was modeled as a sum of production busy time and idle time of Pareto and gamma distributions. *Kuo et al.* (2011) proposed to use neural networks to exploit the production data and tool data of the semiconductor production systems, in order to predict and reduce the production cycle time. *Millstein and Martinich* (2014) developed Takt Time Grouping method to implement kanban-flow manufacturing in a production process with cycle underrun and overrun, where the variations of cycle time were due to manual operation and set-up time randomness. *Kacar et al.* (2016) et al. presented methods in non-integer linear programming to model the cycle time variation in production planning problems. *Larco et al.* (2017) provided methods to estimate the warehouse workers' discomfort and optimize the job assignment, in order to prevent long cycle overrun. *Casalino et al.* (2019) proposed a scheduling method for human-robot collaborative assembly based on the cycle time duration data collected at runtime, adapting to the cycle time underrun and overrun of manufacturing processes. *Ben-Ammar et al.* (2020) studied the integrated production planning and quality control strategies for serial production systems with machines having variable probability distributions of the cycle time. *Roshani et al.* (2020) proposed a hybrid adaptive neighborhood search approach to minimize cycle time variability in multi-sided assembly lines. *Touzani* *et al.* (2021) proposed methods for multi-robot task sequencing and automatic path planning to reduce cycle times of automotive production lines.

1.4 Dissertation Outline

The rest of this dissertation is organized as follows: Chapter 2 presents the (α, β) precision theory for the estimates of $MTBF$ and $MTTR$, along with a probabilistic upper bound of the observation time. Chapter 3 develops the evaluation of the (α_X, β_X)-precise estimate of machine efficiency, throughput, lead time, and work-in-process. The theory concerning the (α_q, β_q)- and $(\alpha_{TP_q}, \beta_{TP_q})$-precise estimate of machine quality parameter q and quality parts throughput TP_q is presented in Chapter 4. Chapter 5 presents the comparison of the (α, β)-Precision Thoery with Markov inequality, Chebyshev inequality and simulations. The modeling, analysis, improvement, and bottleneck identification of production systems with cycle overrun are included in Chapter 6, along with a case study based on an automotive transmission machining line. Conclusion and future works are included in Chapter 7. Proofs, justifications, and relevant simulation results are included in Appendices A-E.

Results presented in Chapter 2 to Chapter 4 of this dissertation have been published as technical report, conference proceeding, and peer-reviewed journal papers, including *Alavian et al.* (2018, 2019); *Alavian et al.* (2021); *Alavian et al.* (2021)[1]. Results of Chapter 6 are reported in *Eun et al.* (2021)[1].

[1]Following a long-standing tradition of Prof. Meerkov's research group, the authors are ordered alphabetically. Kang Liu is identified as the leading author of these publications.

CHAPTER 2

The (α, β)-Precise Estimate of $MTBF$ and $MTTR$: Definition, Calculation, and Observation Time

2.1 Problem Motivation

The Mean Time Between Failures ($MTBF$) and the Mean Time To Repair ($MTTR$) of manufacturing equipment are used in every quantitative method for production systems performance analysis, continuous improvement, and design. To evaluate $MTBF$ and $MTTR$ on the factory floor, random realizations of machine up- and downtime must be measured and then averaged to obtain the estimates, $\widehat{MTBF}$ and $\widehat{MTTR}$. The law of large numbers guarantees that a sufficiently large number of measurements leads to sufficiently precise estimates. However, collecting a large number of measurements requires a long observation time. This is acceptable for off-line evaluation of $MTBF$ and $MTTR$, where historical data could be used.

For on-line applications (e.g., those in the framework of Industry 4.0 and Smart

Manufacturing (*Kagermann et al.* 2013; *Liao et al.* 2017)), the observation period must be as short as possible in order to capture the data necessary for real-time utilization. Remarkably, the question of determining the smallest number of random variable realizations necessary and sufficient for evaluating its average value with the desired accuracy has not been addressed in the literature. In fact, we were able to identify only two papers mentioning this issue. The first one, reporting on Ford's experience (see *Williams* (1994)), lists questions to be asked before $MTTR$ can be evaluated. The second, based on GM's research (see *Inman* (1999)), mentions the number of up- and downtime occurrences, which has been used to estimate up- and downtime probability distributions, without going into specifics of why one or another number has been selected.

This chapter is interned to provide guidance for selecting the smallest number of measurements necessary and sufficient for calculating reliable estimates of $MTBF$ and $MTTR$. The term "reliable" is used to indicate an estimate, which has the desired accuracy with the desired probability. Denoting the accuracy by α and the probability by β (see Section 2.2 for precise formalization), the goal of this chapter is two-fold:

- For a given pair (α, β), calculate how many realizations of machine up- and downtimes are necessary and sufficient to obtain (α, β)-precise estimates of $MTBF$ and $MTTR$.

- Provide a characterization of the observation time required to collect the number of measurements defined by (α, β). This characterization will define temporal properties of $MTBF$ and $MTTR$ evaluation in real-time.

Accordingly, the outline of this chapter is as follows: Section 2.2 presents the definition of (α, β)-precise estimates of $MTBF$ and $MTTR$, along with rigorous formulation of the two problems mentioned above. In Section 2.3, a method for calculating the smallest number of up- and downtime measurements necessary and sufficient for the desired precision of $MTBF$ and $MTTR$ estimates is developed. The issue of observation time is discussed in Section 2.4. Finally, summary of the results obtained is given in Section 2.5. The proofs and justifications are included in Appendices A.

2.2 Definitions and Problems Formulation

Consider an unreliable machine with up- and downtime being random variables with expected values T_{up} and T_{down}, respectively. Obviously, T_{up} and T_{down} are the exact values of $MTBF$ and $MTTR$; we use these two types of notations interchangeably − depending on the issue at hand.

Let $t_{up,i}$ and $t_{down,i}$ be the durations of the i-th occurrence (realization) of up- and downtime, $i = 1, 2, \dots$, respectively. Then, the estimates of $MTBF$ and $MTTR$, based on n observations, are the following random variables:

$$\widehat{T}_{up}(n) := \frac{\sum_{i=1}^{n} t_{up,i}}{n}, \quad \widehat{T}_{down}(n) := \frac{\sum_{i=1}^{n} t_{down,i}}{n}. \tag{2.1}$$

Definition 2.1. The estimates $\widehat{T}_{up}(n)$ and $\widehat{T}_{down}(n)$ are referred to as (α, β)-*precise* if

$$
\begin{aligned}
&P\left\{\frac{|T_{up}-\widehat{T}_{up}(n)|}{T_{up}} \leqslant \alpha\right\} \geqslant \beta, \\
&P\left\{\frac{|T_{down}-\widehat{T}_{down}(n)|}{T_{down}} \leqslant \alpha\right\} \geqslant \beta,
\end{aligned}
\tag{2.2}
$$

or, equivalently,

$$
\begin{aligned}
&P\left\{(1-\alpha)T_{up} \leqslant \widehat{T}_{up}(n) \leqslant (1+\alpha)T_{up}\right\} \geqslant \beta, \\
&P\left\{(1-\alpha)T_{down} \leqslant \widehat{T}_{down}(n) \leqslant (1+\alpha)T_{down}\right\} \geqslant \beta.
\end{aligned}
\tag{2.3}
$$

Clearly, this definition implies that the accuracy of the estimates is quantified by α and their likelihood by β. For instance, if $\alpha = 0.05$ and $\beta = 0.9$, the appropriately selected value of n guarantees that $\widehat{T}_{up}(n)$ and $\widehat{T}_{down}(n)$ are within $\pm 5\%$ of T_{up} and T_{down}, respectively, and this event takes place with probability at least 0.9.

Definition 2.2. The smallest $n_T^*(\alpha, \beta)$, which guarantees (2.2), is referred to as the *critical number of measurements.*

The first problem addressed in this chapter consists of two parts:

Problem 1a: For a given pair (α, β), calculate $n_T^*(\alpha, \beta)$ for machines with exponential reliability model (i.e., with up- and downtime distributed exponentially with parameters λ and μ, respectively).

Problem 1b: Generalize the results of Problem 1a to machines with non-exponential reliability models, having the coefficient of variation, CV, less than 1. Note that, as it is shown in *Li and Meerkov* (2009), if the machine breakdown rate (respectively, repair rate) is an increasing function of time, the resulting distribution of uptime (respectively, downtime) has $CV < 1$. Empirical evidence that manufacturing equipment on the factory floor pratically always has $CV < 1$ can be found in *Inman* (1999).

The second problem consists of evaluating the observation time (OT) necessary to collect $n_T^*(\alpha, \beta)$ realizations of $t_{up,i}$ and $t_{down,i}$. If T_{up} and T_{down} were known, the mean observation time (MOT) would be

$$MOT = n_T^*(\alpha, \beta)(T_{up} + T_{down}). \tag{2.4}$$

Since T_{up} and T_{down} are unknown, the approach employed here is based on using $n_{T,0} < n_T^*(\alpha, \beta)$ initial measurements to calculate the estimates $\widehat{T}_{up}(n_{T,0})$ and $\widehat{T}_{down}(n_{T,0})$ and then defining an estimate of the remaining observation time ($\widehat{ROT}$) as the following random variable:

$$\widehat{ROT}(n_{T,0}) = (n_T^*(\alpha, \beta) - n_{T,0}) \left(\widehat{T}_{up}(n_{T,0}) + \widehat{T}_{down}(n_{T,0})\right). \tag{2.5}$$

In reality, however, the remaining observation time is a random variable given by

$$ROT(n_{T,0}) = \sum_{i=n_{T,0}+1}^{n_T^*(\alpha,\beta)} (t_{up,i} + t_{down,i}). \tag{2.6}$$

The relationship between these two random variables can be characterized by the following inequality:

$$P\left\{ROT(n_{T,0}) < a\widehat{ROT}(n_{T,0})\right\} \geqslant b, \tag{2.7}$$

where a is referred to as a safety factor and b is the desired probability.

Definition 2.3. The smallest integer $n^*_{T,0}(a,b) < n^*_T(\alpha,\beta)$, which guarantees (2.7), is referred to as the *critical number of initial measurements.*

Problem 2: For a given pair (a,b), calculate $n^*_0(a,b)$ for machines with exponential reliability model.

These two problems are solved in Sections 2.3 and 2.4, respectively.

2.3 Evaluating Critical Number of Measurements

2.3.1 Exponential Machines

2.3.1.1 Exact value of $n^*_T(\alpha,\beta)$

Theorem 2.1. *The critical number, $n^*_T(\alpha,\beta)$, for the case of machines with exponential reliability model is the smallest integer n, which satisfies the following inequality:*

$$\beta \leqslant \sum_{i=0}^{n-1} \frac{1}{i!} e^{-(1-\alpha)n} \left((1-\alpha)n\right)^i - \sum_{i=0}^{n-1} \frac{1}{i!} e^{-(1+\alpha)n} \left((1+\alpha)n\right)^i . \tag{2.8}$$

Proof. See Appendix A. □

Corollary 2.2. *The critical number $n^*_T(\alpha,\beta)$ is the same for both $MTBF$ and $MTTR$.*

Proof. Follows immediately from the fact that the right-hand side of (2.8) is independent of the parameter of the exponential distribution involved. □

Clearly, this corollary is of substantial practical importance. It implies that the smallest number of measurements necessary and sufficient to identify reliability characteristics of a single machine or multiple machines in a production system are the same.

The value of $n_T^*(\alpha,\beta)$ can be obtained by monotonically increasing n in (2.8) until the inequality is satisfied. Based on this calculation, the behavior of $n_T^*(\alpha,\beta)$ is illustrated in Fig. 2.1. As expected, this function is monotonically increasing in β and monotonically decreasing in α.

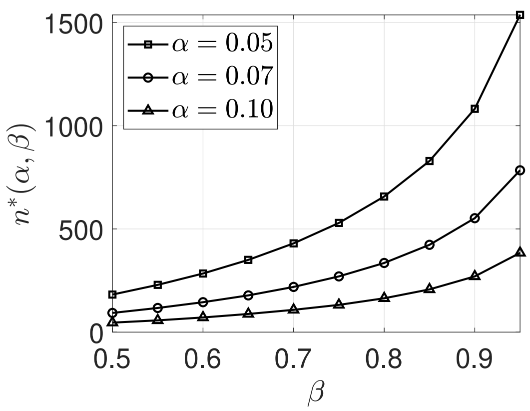


Figure 2.1. Critical number n_T^* as a function of β and α

2.3.1.2 Gaussian approximation of $n_T^*(\alpha,\beta)$

Along with (2.8), it is desirable to have an analytical expression for $n_T^*(\alpha,\beta)$. Such an expression can be derived using the fact that, while $\widehat{T}_{up}(n_T^*)$ and $\widehat{T}_{down}(n_T^*)$ are Erlang random variables, for sufficiently large n_T^* and under appropriate normalization, they are close to Gaussian random variable $\mathcal{N}(0,1)$. Based on this approximation, the following is obtained:

Proposition 2.3. *The Gaussian approximation of the critical number, $n^*_{T,G}(\alpha,\beta)$, is given by:*

$$n^*_{T,G}(\alpha,\beta) = \left\lceil 2\left(\frac{\mathrm{erf}^{-1}(\beta)}{\alpha}\right)^2 \right\rceil, \qquad (2.9)$$

where the ceiling operator $\lceil x \rceil$ denotes the smallest integer larger than x and $\mathrm{erf}^{-1}(y)$ *is the inverse of the error function,* $\mathrm{erf}(y) = \frac{1}{\sqrt{\pi}} \int_{-y}^{y} e^{-t^2} dt$.

Justification. See Appendix A. □

The accuracy of Gaussian approximation is illustrated by comparison of $n^*_T(\alpha,\beta)$ and $n^*_{T,G}(\alpha,\beta)$ given in Tables 2.1 and 2.2. As one can see, for all values of α and β of practical importance analyzed, $n^*_T(\alpha,\beta)$ and $n^*_{T,G}(\alpha,\beta)$ are almost the same.

Table 2.1. Critical number $n^*_T(\alpha,\beta)$

α \ β	**0.7**	**0.8**	**0.85**	**0.9**	**0.95**
0.02	2686	4106	5181	6764	9604
0.04	672	1026	1295	1691	2401
0.06	299	456	576	751	1067
0.08	168	257	324	423	600
0.10	108	164	207	270	384
0.20	27	41	52	67	96

The calculation of $n^*_{T,G}(\alpha,\beta)$ is orders of magnitude faster than that of $n^*_T(\alpha,\beta)$, which allows for more detailed investigation of the critical number n^*. For instance, Fig. 2.2 presents the contour plot of critical number $n^*_{T,G}(\alpha,\beta)$, calculated using (2.9). This plot offers guidance for selecting $n^*_{T,G}$ for the desired α and β. Indeed, if $\alpha = 0.05$ and $\beta = 0.9$, from Fig. 2.2 we obtain $n^*_{T,G}(\alpha,\beta) \approx 1000$. On the other

Table 2.2. Critical number $n^*_{T,G}(\alpha, \beta)$

α \ β	**0.7**	**0.8**	**0.85**	**0.9**	**0.95**
0.02	2686	4106	5181	6764	9604
0.04	672	1027	1296	1691	2401
0.06	299	457	576	752	1068
0.08	168	257	324	423	601
0.10	108	165	208	271	385
0.20	27	42	52	68	97

hand, if $\alpha = 0.15$ and $\beta = 0.75$, $n^*_{T,G}(\alpha, \beta) \approx 60$. In some cases, the number of measurements that one can collect on the factory floor during two weeks of observation period is between 50 and 100. Thus, in these situations only relatively inaccurate estimates $\widehat{T}_{up}(n^*_T)$ and $\widehat{T}_{down}(n^*_T)$ could be obtained.

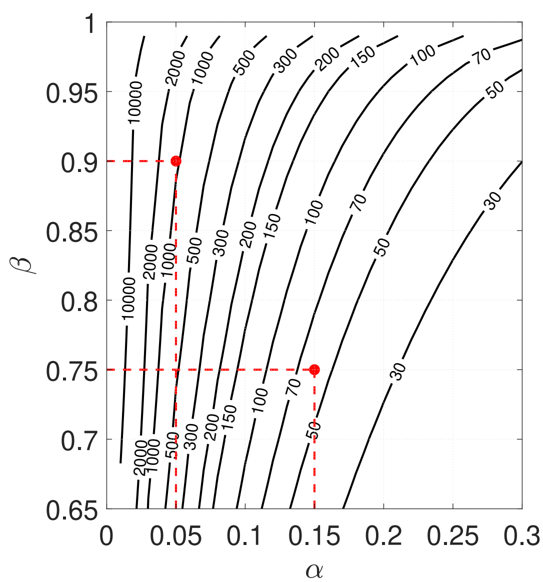


Figure 2.2. Contour plot of $n^*_{T,G}(\alpha, \beta)$

2.3.2 Non-exponential Machines

To investigate the critical number of measurements in the non-exponential case, we consider machines obeying Weibull, gamma, and log-normal reliability models with $MTBF = 10$ and $CV \in \{0.1, 0.25, 0.5, 0.75\}$ and evaluate by simulations $n^*_{T,\text{non-exp}}(\alpha, \beta; CV)$ for $\alpha = 0.05$ and $\beta \in \{0.65, 0.70, \ldots, 0.95\}$. The results are summarized in Fig. 2.3, where $n^*_T(\alpha, \beta)$ for exponential distribution is shown for comparison. From this figure, we conclude:

Observation 2.1. For all non-exponential machines analyzed:

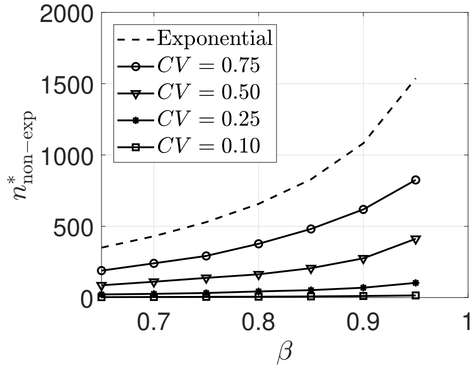
- $n^*_{T,\text{non-exp}}(\alpha, \beta; CV) < n^*_T(\alpha, \beta)$;
- $n^*_{T,\text{non-exp}}(\alpha, \beta; CV)$ approaches $n^*_T(\alpha, \beta)$ when $CV \to 1$;
- $n^*_{T,\text{non-exp}}(\alpha, \beta; CV)$ is practically independent of the machine up- and downtime distribution as long as CV is the same.

Thus, the number of measurements, selected under the exponential assumption, can be used as an upper bound for non-exponential machines, provided $CV < 1$.

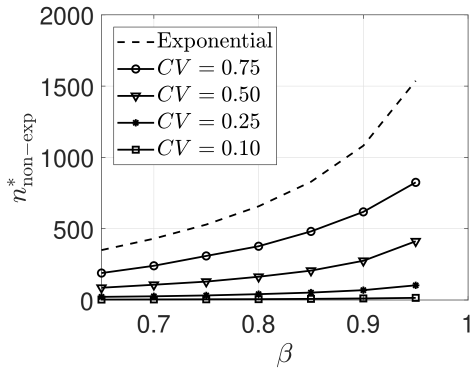
We hypothesize that Observation 2.1 holds not only for the distributions analyzed, but for any unimodal distribution of up- and downtime with $CV < 1$.

2.4 Evaluating Critical Number of Initial Measurements

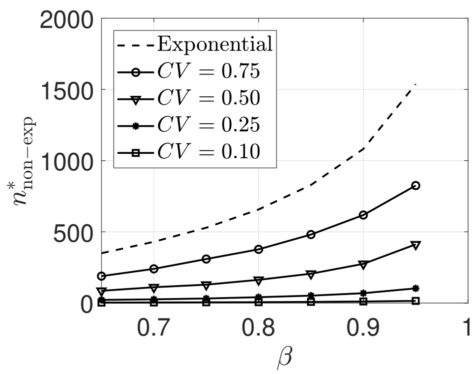
The total observation time (TOT) to collect $n^*(\alpha, \beta)$ measurements of up- and downtime can be represented as a sum of two random variables – one representing



(a) Weibull reliability model



(b) gamma reliability model



(c) log-normal reliability model

Figure 2.3. Critical numbers $n^*_{\text{non-exp}}$ and n^*_T for $\alpha = 0.05$ as functions of β and CV

the initial observation time (IOT), i.e., the time before initial $n_{T,0} < n_T^*(\alpha,\beta)$ measurements have been collected, and the other representing the remaining observation time (ROT), i.e., the time to collect the remaining $[n_T^*(\alpha,\beta) - n_{T,0}]$ measurements:

$$
\begin{aligned}
TOT &= IOT + ROT \qquad (2.10)\\
&= \sum_{i=1}^{n_{T,0}} (t_{up,i} + t_{down,i}) + \sum_{i=n_{T,0}+1}^{n_T^*(\alpha,\beta)} (t_{up,i} + t_{down,i}). \qquad (2.11)
\end{aligned}
$$

After the initial $n_{T,0}$ measurements have been collected, the first term in this sum becomes a realization of IOT, i.e., a constant denoted as $IOT(n_{T,0})$. The second term, as discussed in Section 2.2, can be approximated by the following random variable:

$$
\widehat{ROT}(n_{T,0}) = (n_T^*(\alpha,\beta) - n_{T,0}) \left(\widehat{T}_{up}(n_{T,0}) + \widehat{T}_{down}(n_{T,0}) \right), \qquad (2.12)
$$

where $\widehat{T}_{up}(n_{T,0})$ and $\widehat{T}_{down}(n_{T,0})$ are the estimates of $MTBF$ and $MTTR$ based on $n_{T,0}$ initial measurements. This implies that TOT can be characterized by the following stochastic upper bound:

$$
P\left\{ IOT(n_{T,0}) + ROT(n_{T,0}) < IOT(n_{T,0}) + a\widehat{ROT}(n_{T,0}) \right\} \geqslant b \qquad (2.13)
$$

or, equivalently (as in (13)),

$$
P\left\{ ROT(n_{T,0}) < a\widehat{ROT}(n_{T,0}) \right\} \geqslant b, \qquad (2.14)
$$

where a is a safety factor and b the desired probability. In Section II, the smallest $n_{T,0}$ satisfying this inequality has been referred to as the critical number of initial measurements, $n^*_{T,0}(a,b)$. When $n^*_{T,0}(a,b)$ is determined, $\widehat{ROT}(n_{T,0})$ can be calculated using (2.12) and then the upper bound of TOT evaluated using (2.13). A method for calculating $n^*_{T,0}(a,b)$ is provided by

Proposition 2.4. *Under Gaussian assumption, for any* (a,b) *with* $a > 1$, $0.5 < b < 1$, *and* $n^*_{T,G}(\alpha,\beta)$ *satisfying*

$$n^*_{T,G}(\alpha,\beta) \geqslant \left\lceil \frac{(a+1)^2}{(a-1)^2} \cdot 2[\mathrm{erf}^{-1}(2b-1)]^2 \right\rceil, \tag{2.15}$$

the critical number of initial measurements, $n^*_{T,0,G}(a,b)$, *is:*

$$n^*_{0,G}(a,b;n^*_{T,G}) = \left\lceil \frac{Bn^*_{T,G} + a^2 - 1 - \sqrt{(Bn^*_{T,G} + a^2 - 1)^2 - 4Bn^*_{T,G}a^2}}{2B} \right\rceil, \tag{2.16}$$

where $n^*_{T,G} = n^*_{T,G}(\alpha,\beta)$ *and*

$$B = \frac{(a-1)^2}{2[\mathrm{erf}^{-1}(2b-1)]}. \tag{2.17}$$

Justification. See Appendix A. □

The contour plots of $n^*_{T,0,G}$ as a function of a and b are shown in Fig. 2.4 for $n^*_{T,G}(\alpha = 0.1, \beta = 0.9) = 270$ and $n^*_{T,G}(\alpha = 0.05, \beta = 0.9) = 1027$. The greyed parts in Fig. 2.4 indicate the areas, where inequality (2.15) does not hold. These contour plots indicate the following:

- For $a \leqslant 1.1$, $n^*_{T,0,G}(a, b; n^*_{T,G})$ depends significantly on $n^*_{T,G}$. For instance, if $n^*_{T,G} = 270$, no $n^*_{T,0,G}(a, 0.9)$ satisfying (2.15) exists, whereas if $n^*_{T,G} = 1027$, it does: $n^*_{T,0,G}(1.1, 0.9) \approx 250$ guarantees that (2.13) provides a relatively tight bound of TOT (having $a = 1.1$) and, this bound takes place with a relatively large probability (having $b = 0.9$).

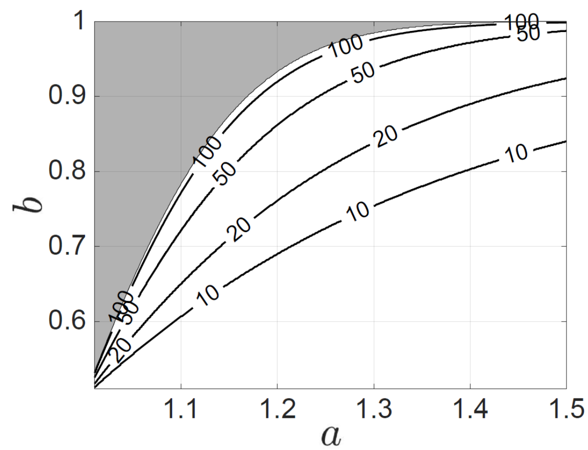
- For $a \geqslant 1.3$, $n^*_{T,0,G}(a, b; n^*_{T,G})$ is practically independent of $n^*_{T,G}$. For instance, if $a = 1.3$ and $b = 0.9$, $n^*_{T,0,G}$ is approximately 33 for both $n^*_{T,G} = 270$ and $n^*_{T,G} = 1027$. Thus, if in a particular application $a = 1.3$ is viewed an acceptable safety factor, $IOT(n^*_{T,0,G}) + 1.3\widehat{ROT}(n^*_{T,0,G})$, where $n^*_{T,0,G} = 33$, provides an estimate of TOT taking place with probability 0.9 for both $(\alpha, \beta) = (0.1, 0, 9)$ and for $(\alpha, \beta) = (0.05, 0, 9)$.

2.5 Summary

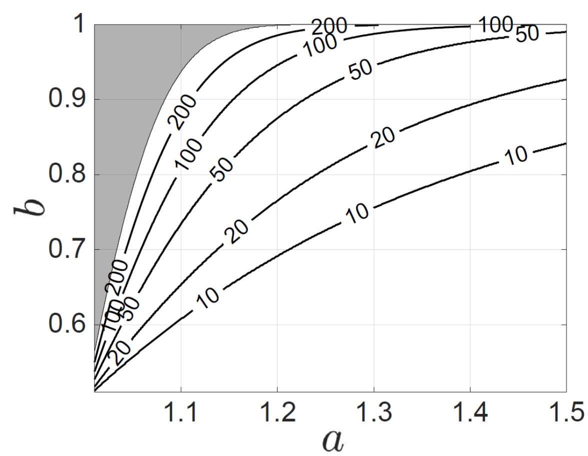
This chapter presents the following results:

- A method for calculating the smallest number, $n^*_T(\alpha, \beta)$, of up- and downtime measurements, necessary and sufficient to obtain (α, β)-precise estimates of $MTBF$ and $MTTR$ for machines with exponential reliability model, where α represents the estimate's accuracy and β its probability.

  - Since this method is combinatorial in nature and the resulting calculations are based on iterations, this chapter derives an analytical, Gaussian approximation for calculating $n^*_T(\alpha, \beta)$ and shows that the resulting



(a) $n^*_{T,G}(\alpha, \beta) = 270$



(b) $n^*_{T,G}(\alpha, \beta) = 1027$

Figure 2.4. Contour plots of $n^*_{T,0,G}(a, b; n^*_T)$

$n^*_{T,G}(\alpha, \beta)$ is almost the same as $n^*_T(\alpha, \beta)$ and the calculation of the former is orders of magnitude faster than the latter.

- ○ For non-exponential machines, it is shown, by simulations, that $n^*_{T,G}(\alpha, \beta)$ is the upper bound of the number of measurements required for evaluating (α, β)-precise estimates of $MTBF$ and $MTTR$ in machines with Weibull, gamma, and log-normal reliability models with $CV < 1$.

- A method for calculating an upper bound of the observation time required to collect $n^*_T(\alpha, \beta)$ measurements of machine up- and downtime. The approach is based on calculating the smallest number of initial measurements, $n^*_{T,0}(a, b)$ (where $a > 1$ represents the safety factor and b the desired probability that the upper bound indeed takes place) and using the resulting estimates of $MTBF$ and $MTTR$ to evaluate the remaining observation time.

From these results follows a number of quantitative and qualitative conclusions. The quantitative conclusion is: For practically important α and β, the value of n^*_T is quite large. For example, if $\alpha = 0.05$ and $\beta = 0.9$, then $n^*_T(\alpha, \beta) \approx 1000$. This implies that if, for instance, a machine has about 100 of up and down events per week, the observation period would be about ten weeks, which is hardly acceptable in practice due to the natural machine efficiency degradation and due to "low frequency" of $MTBF$ and $MTTR$ estimates to be used for a "higher frequency" decision-making.

The main qualitative conclusion is: The smallest number of up- and downtime measurements necessary and sufficient for evaluating $MTBF$ and $MTTR$ in exponential machines is independent of the exponential distribution parameter. This implies that the same number of measurements is required for identifying parameters of all machines comprising a production system (under the exponential assumption).

CHAPTER 3

The (α_X, β_X)-Precise Estimate of Machine Efficiency, Throughput, Lead Time, and Work-In-Process

3.1 Problem Motivation

Estimates of production system performance metrics, such as machine efficiency (e), system throughput (TP), production lead time (LT), and work-in-process (WIP) are necessary for managing production systems and for evaluating effectiveness of potential system modifications leading to the desired productivity improvement. The calculation of these estimates requires machine reliability characteristics, primarily Mean Time Between Failures ($MTBF$) and Mean Time To Repair ($MTTR$), which can be obtained using factory floor measurements of machines' up- and downtime realizations. In this regard a question arises: What is the smallest number of measurements required to ensure desired accuracy of the induced estimates $\widehat{e}$, $\widehat{TP}$, $\widehat{LT}$, and $\widehat{WIP}$?

A similar question, concerning $MTBF$ and $MTTR$, has been posed and answered in Chapter 2. Since these estimates are random variables, the estimate of machine efficiency, e, calculated according to

$$\widehat{e} = \frac{\widehat{T}_{up}}{\widehat{T}_{up} + \widehat{T}_{down}}, \tag{3.1}$$

is also random. Similarly, the estimates $\widehat{TP}$, $\widehat{LT}$, and $\widehat{WIP}$ are random variables too. Based on $\widehat{T}_{up}$ and $\widehat{T}_{down}$, they can be calculated using a number of production systems performance analysis techniques developed in *Viswanadham and Narahari* (1992), *Askin and Standridge* (1993), *Buzacott and Shanthikumar* (1993), *Papadopolous et al.* (1993), *Gershwin* (1994), *Altiok* (1997), *Li and Meerkov* (2009), *Papadopoulos et al.* (2009) and *Curry and Feldman* (2010). In the current chapter, we use the aggregation technique of *Li and Meerkov* (2009), primarily because it provides proofs of convergence of the recursive iteration procedures arising in analysis of systems with more than two machines and addresses, in a unified framework, all the performance metrics mentioned above.

Quantifying the induced precision of $\widehat{e}$, $\widehat{TP}$, $\widehat{LT}$ and $\widehat{WIP}$ by (α_X, β_X), where $X \in \{e, TP, LT, WIP\}$, this chapter provides an answer to the question posed above in terms of serial production lines with machines obeying the exponential reliability model. This is accomplished by calculating (α_X, β_X), induced by (α, β), and solving the inverse problem, i.e., calculating (α, β) for $\widehat{T}_{up}$ and $\widehat{T}_{down}$ leading to the desired (α_X, β_X), $X \in \{e, TP, LT, WIP\}$.

The issues, addressed in this chapter, are of importance for three reasons. The first one is fairly straightforward: Since all methods of performance analysis, design, and continuous improvement of production systems assume that $MTBF$ and $MTTR$ are known exactly, it is important to know what, in fact, is the accuracy of the calculated performance metrics, taking into account that merely estimates of $MTBF$ and $MTTR$ are available.

The second reason is of practical importance and consists of the following: It has been shown in Chapter 2 that to obtain relatively precise estimates of $MTBF$ and $MTTR$, the smallest number of up- and downtime measurements may be quite large. For instance, if $(\alpha, \beta) = (0.05, 0.9)$, this number, denoted as $n_T^*(\alpha, \beta)$, is 1083. Collecting these measurements may require a long observation period, which might be unacceptable in real-time applications, e.g., in the framework of Smart Manufacturing as a part of Industry 4.0 *Kagermann et al.* (2013); *Liao et al.* (2017). If, however, it turns out that the induced $\alpha_X < \alpha$ and/or $\beta_X > \beta$, a sufficiently precise estimates of X would require a smaller number of measurements than equally precise estimates of $MTBF$ and $MTTR$.

Finally, the third reason, also of industrial significance, is as follows: In most analysis and continuous improvement projects, it is obvious what precision of $\widehat{X} \in \{\widehat{e}, \widehat{TP}, \widehat{LT}, \widehat{WIP}\}$ is required for the problem at hand (e.g., $\alpha_X = 0.05$ and $\beta_X = 0.9$). However, it is difficult to predict which precision of $MTBF$ and $MTTR$ would be necessary to guarantee the required precision of $\widehat{X}$. Therefore, evaluating (α, β), based on the required (α_X, β_X), would guide the production managers and engineering/reserach personnel in the problem of selecting the smallest number of measurements for evaluating $\widehat{X}$ with the desired precision.

The novel quantitative results obtained in this chapter are:

- Analytical expressions for (α_X, β_X), $X \in \{e, TP, LT, WIP\}$, as functions of (α, β)-precise estimates of T_{up} and T_{down}.

- Analytical expressions for the number of machines' up- and downtime measurements, required to obtain the desired (α_X, β_X)-precise estimates of $X \in \{e, TP, LT, WIP\}$.

In addition to quantitative results, the chapter provides qualitative insights into variability properties of production system performance metrics. To describe them, let $n^*_T(\gamma, \delta)$ denote the smallest number of up- and downtime measurements required to obtain (γ, δ)-precise estimates of T_{up} and T_{down}, while $n^{**}_X(\gamma, \delta)$ denotes the smallest number of up- and downtime measurements required to obtain an equally precise estimate of $X \in \{e, TP, LT, WIP\}$.

Definition 3.1. Performance metric $X \in \{e, TP, LT, WIP\}$ is:

- ○ *variability contracting* if $n^{**}_X(\gamma, \delta) < n^*_T(\gamma, \delta)$;

- ○ *variability expanding* if $n^{**}_X(\gamma, \delta) > n^*_T(\gamma, \delta)$.

In practical terms, this definition implies that X is variability contracting, if its sufficiently precise estimate $\widehat{X}$ can be obtained using less precise estimates of its arguments. In contrast, variability expanding X implies that the precision of $\widehat{X}$ is lower than that of its arguments under the same number of measurements.

In terms of these concepts, this chapter shows that under some practice-inspired conditions,

- e and TP are variability contacting;
- LT, and WIP are variability expanding.

The methods developed in this chapter are intended, primarily, for production systems managerial/engineering and research personnel involved in either daily decision-making or evaluating efficacy of potential improvement projects. In practice, managers and industrial engineers often use a rule-of-thumb, which assumes that about two-week worth of data on machines' up- and downtime measurements are sufficient to evaluate efficacy of potential improvement projects. The current chapter shows that this may or may not be the case – either more or less measurements might be necessary, depending on the performance metrics addressed and the precision sought.

The outline of this chapter is as follows: Section 3.2 provides formulation of the problems addressed. Sections 3.3-3.6 analyze induced (α_X, β_X)-precise estimates of $X \in \{e, TP, LT, WIP\}$ and solve the corresponding inverse problems. Section 3.7 offers a summary of analytical expressions derived in this work and provides a comparative illustration of the number of measurements required for evaluating $\widehat{e}$, $\widehat{TP}$, $\widehat{LT}$, and $\widehat{WIP}$ with a given precision. Finally, Section 3.8 formulates the conclusions and topics for future research. The list of abbreviations and notations is given after the conclusions. All justifications are provided in Appendices B.

3.2 Problems Formulation

Consider a serial production line with M exponential machines and its performance metrics $X \in \{e, TP, LT, WIP\}$. When the exact values of up- and downtime,

$T_{up,i}$ and $T_{down,i}$, $i = 1, \ldots, M$, are available, these performance metrics can be calculated either by a closed formula (for e) or by recursive aggregation procedures developed in *Li and Meerkov* (2009) and its improved version *Bai et al.* (2021) (for TP), and *Meerkov and Yan* (2016) (for LT and WIP). When $T_{up,i}$ and $T_{down,i}$ are not available, their (α, β)-precise estimates, $\widehat{T}_{up,i}\,(n_T^*(\alpha, \beta))$ and $\widehat{T}_{down,i}\,(n_T^*(\alpha, \beta))$ can be used instead, leading to the estimates $\widehat{e}$, $\widehat{TP}$, $\widehat{LT}$, and $\widehat{WIP}$, which are random variables with the precision induced by $\widehat{T}_{up,i}\,(n_T^*(\alpha, \beta))$ and $\widehat{T}_{down,i}\,(n_T^*(\alpha, \beta))$. Similar to (2.2), we quantify the (α_X, β_X)-precision of $\widehat{X} \in \{\widehat{e}, \widehat{TP}, \widehat{LT}, \widehat{WIP}\}$ by:

$$P\left\{\frac{|X - \widehat{X}(n_T^*(\alpha, \beta))|}{X} \leqslant \alpha_X\right\} \geqslant \beta_X. \tag{3.2}$$

Based on (3.2), we introduce the following problems:

Induced precision problem for $\boldsymbol{X}$: Given (α, β)-precise estimates $\widehat{T}_{up,i}\,(n_T^*(\alpha, \beta))$ and $\widehat{T}_{down,i}\,(n_T^*(\alpha, \beta))$, $i = 1, \ldots, M$, calculate the induced precision $(\alpha_X^{ind}, \beta_X^{ind})$ of $\widehat{X} \in \{\widehat{e}, \widehat{TP}, \widehat{LT}, \widehat{WIP}\}$.

Inverse problem for $\boldsymbol{X}$: Given a desired pair (γ, δ), calculate the smallest number of machines' up- and downtime measurements, $n_X^{**}(\gamma, \delta)$, required to obtain (γ, δ)-precise estimate $\widehat{X} \in \{\widehat{e}, \widehat{TP}, \widehat{LT}, \widehat{WIP}\}$.

The induced precision and inverse problems for e, TP, LT, and WIP are considered in Sections 3.3-3.6, respectively.

3.3 Induced Precision and Inverse Problem for Machine Efficiency

3.3.1 Solution of Induced Precision Problem for e

The value of α_e^{ind} is given by:

Proposition 3.1. *Given* $\frac{|Tup - \widehat{T}_{up}(n_T^*(\alpha,\beta))|}{T_{up}} \leqslant \alpha$ *and* $\frac{|T_{down} - \widehat{T}_{down}(n_T^*(\alpha,\beta))|}{T_{down}} \leqslant \alpha$*, the smallest* α_e*, which satisfies* $\frac{|e - \widehat{e}(n_T^*(\alpha,\beta))|}{e} \leqslant \alpha_e$ *with accuracy up to* $O(\alpha^2)$*, is given by*

$$\alpha_e^{ind} = 2\alpha(1 - \widehat{e}(n_T^*(\alpha,\beta))). \tag{3.3}$$

Justification. See Appendix B. □

Expression (3.3) implies the following:

- Since $\widehat{e}\,(n_T^*(\alpha,\beta))$ is a random variable, α_e^{ind} is random as well. However, when the realization of $\widehat{e}\,(n_T^*(\alpha,\beta))$ is calculated according to

$$\widehat{e}(n_T^*(\alpha,\beta)) = \frac{\widehat{T}_{up}(n_T^*(\alpha,\beta))}{\widehat{T}_{up}(n_T^*(\alpha,\beta)) + \widehat{T}_{down}(n_T^*(\alpha,\beta))}, \tag{3.4}$$

  the deterministic number α_e^{ind} quantifies the accuracy of $\widehat{e}$ in the sense of (3.2).

- For all $\widehat{e}\,(n_T^*(\alpha,\beta)) > 0.5$ (which is a practical case), $\alpha_e^{ind} < \alpha$, i.e., the accuracy of $\widehat{e}\,(n_T^*(\alpha,\beta))$ is higher than the accuracy of the underlying $\widehat{T}_{up}(n_T^*(\alpha,\beta))$ and $\widehat{T}_{down}(n_T^*(\alpha,\beta))$. For instance, when $\widehat{e}\,(n_T^*(\alpha,\beta)) = 0.75$, $\alpha_e^{ind} = 0.5\alpha$; when $\widehat{e}\,(n_T^*(\alpha,\beta)) = 0.95$, $\alpha_e^{ind} = 0.1\alpha$.

As far as the induced probability, β_e^{ind}, is concerned, its value is given by:

Theorem 3.2. *If the machine obeys the exponential reliability model and* α_e^{ind} *is selected as* (3.3)*, the resulting* β_e^{ind} *is given by*

$$\begin{aligned}\beta_e^{ind} = \;& \sum_{i=0}^{n^*-1} \frac{(2n^*-2-i)!}{(n^*-1-i)!(n^*-1)!}\big[(1+2\alpha)^{n^*}(2+2\alpha)^{-2n^*+i+1} \\ & -(1-2\alpha)^{n^*}(2-2\alpha)^{-2n^*+i+1}\big],\end{aligned} \tag{3.5}$$

where n^* *denotes* $n^*(\alpha, \beta)$.

Proof. See Appendix B. □

Thus, β_e^{ind} depends explicitly on α, implicitly on β (through $n^*(\alpha, \beta)$), and does not depend on T_{up} and T_{down} and, therefore, on e.

The values of β_e^{ind} for various pairs (α, β) are illustrated in Table 3.1. As one can see, for all (α, β) investigated, $\beta_e^{ind} > \beta$. Thus, α_e^{ind} is smaller than α (if $\widehat{e} > 0.5$) and β_e^{ind} is larger than β . In other words, the induced (α_e, β_e)-precise estimate of e is better than (α, β)-precise estimates of T_{up} and T_{down}.

Table 3.1. Values of β_e^{ind} as a function of α and β

α \ β	**0.7**	**0.8**	**0.85**	**0.9**	**0.95**
0.02	0.8573	0.9300	0.9582	0.9799	0.9944
0.04	0.8575	0.9299	0.9580	0.9797	0.9942
0.06	0.8578	0.9298	0.9576	0.9794	0.9940
0.08	0.8576	0.9294	0.9571	0.9788	0.9937
0.10	0.8585	0.9294	0.9568	0.9782	0.9933

As a numerical example, assume $\alpha = 0.1$, $\beta = 0.9$, and $\widehat{e}(n^*(\alpha, \beta)) = 0.8$. Then, according to (3.3) and (3.5), $\alpha_e^{ind} = 0.044$ and $\beta_e^{ind} = 0.9782$. Thus, the structure of (3.1) induces a significantly more precise estimate of e than that of T_{up} and T_{down}.

The Gaussian approximation of β_e^{ind} is given by:

Proposition 3.3. *The Gaussian approximation of β_e^{ind} is given by*

$$\beta_e^{ind} = \text{erf}\left(\alpha\sqrt{n_T^*(\alpha,\beta)}\right), \tag{3.6}$$

where $n_T^*(\alpha,\beta)$ is defined by (2.9)*.*

Justification. See Appendix B. □

A comparison of β_e^{ind} and $\beta_{e,G}^{ind}$ is given in Tables 3.1 and 3.2. As one can see, these values are almost always the same.

Table 3.2. Values of $\beta_{e,G}^{ind}$ as a function of α and β

α \ β	**0.7**	**0.8**	**0.85**	**0.9**	**0.95**
0.02	0.8573	0.9301	0.9582	0.9800	0.9944
0.04	0.8575	0.9300	0.9582	0.9800	0.9944
0.06	0.8577	0.9300	0.9583	0.9799	0.9944
0.08	0.8575	0.9303	0.9583	0.9800	0.9944
0.10	0.8584	0.9299	0.9581	0.9799	0.9944

Omitting the ceiling operator in (2.9) and substituting it in (3.6), we obtain:

$$\beta_e^{ind} = \text{erf}\left(\sqrt{2}\text{erf}^{-1}(\beta)\right). \tag{3.7}$$

This expression shows that β_e^{ind} is independent of α_e^{ind} and, in addition, quantifies to which extent β_e^{ind} is larger than β. For instance, if $\beta = 0.5$, the value of β_e^{ind} is larger than 0.65; if $\beta = 0.75$, β_e^{ind} is about 0.9.

Thus, under the assumption that $\widehat{e}(n_T^*(\alpha, \beta)) > 0.5$, it follows from (3.3) and (3.7) that the induced precision of $\widehat{e}$ is higher than the precision of the underlining $\widehat{T}_{up}$ and $\widehat{T}_{down}$.

3.3.2 Solution of Inverse Problem for *e*

Proposition 3.4. *For a given (γ, δ), the critical number of up- and downtime measurements to ensure (γ, δ)-precise estimate of e is given by:*

$$n_e^{**}(\gamma, \delta) = \left\lceil \left(\frac{2(1 - \widehat{e})\mathrm{erf}^{-1}(\delta)}{\gamma} \right)^2 \right\rceil, \tag{3.8}$$

where $\widehat{e} = \widehat{e}(n_e^{**}(\gamma, \delta))$.

Justification. See Appendix B. □

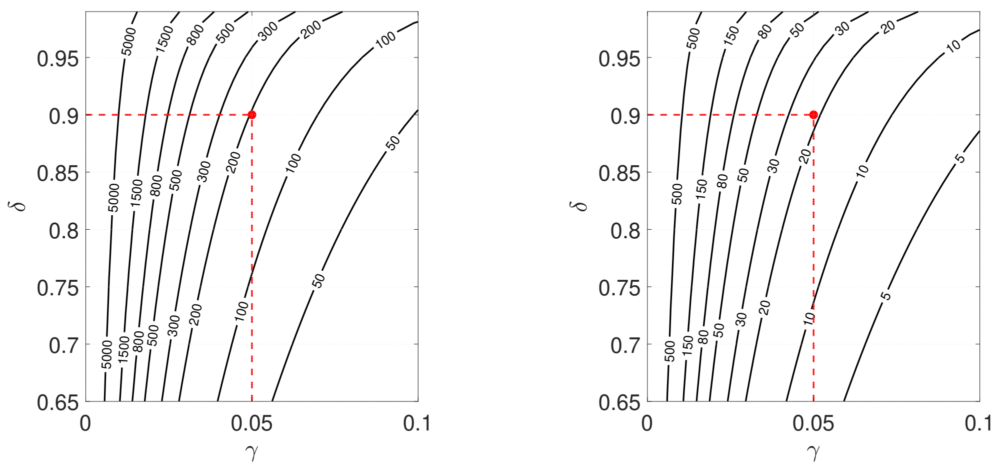
As one can see, $n_e^{**}(\gamma, \delta)$ turns out to be dependent on $\widehat{e}$ and, thus, can be denoted as $n_e^{**}(\gamma, \delta; \widehat{e})$. The contour plots of $n_e^{**}(\gamma, \delta; \widehat{e})$ are shown in Figure 3.1 for $\widehat{e} = 0.7$ and $\widehat{e} = 0.9$. These plots and the plot of Figure 2.2 allow one to compare $n_T^*(\gamma, \delta)$ with $n_e^{**}(\gamma, \delta; \widehat{e})$ for various values of $\widehat{e}$. Indeed, for $\gamma = 0.05$ and $\delta = 0.9$,

- $n_T^*(\gamma, \delta) \approx 1000$;
- $n_e^{**}(\gamma, \delta; \widehat{e}) \approx 200$ if $\widehat{e} = 0.7$ and $n_e^{**}(\gamma, \delta; \widehat{e}) \approx 20$ if $\widehat{e} = 0.9$.

This leads to:

Observation 3.1. Comparing $n_T^*(\gamma, \delta)$, defined by (2.9), with $n_e^{**}(\gamma, \delta)$, defined by (3.8), we conclude that for $\widehat{e} > 0.5$ performance metric e is variability contracting.

Note that $\widehat{e} > 0.5$ is a sufficient condition. The necessary and sufficient condition is $\widehat{e} > 0.2929$. We use, however, $\widehat{e} > 0.5$ in order to maintain that $\alpha_e^{ind} > \alpha$ (see (3.3)).



(a) $\widehat{e} = 0.7$ (b) $\widehat{e} = 0.9$

Figure 3.1. Contour plots of $n_e^{**}(\gamma, \delta; \widehat{e})$

Concluding this subsection, an issue of practical importance must be addressed: The right-hand side of (3.8) includes the machine efficiency estimate $\widehat{e}$ calculated based on $n_e^{**}(\gamma, \delta)$, which is unknown. To alleviate this problem, $\widehat{e}(n_e^{**}(\gamma, \delta))$ is approximated and subsequently used in the right-hand side of (3.8). Since a similar approximation is necessary for other performance metric, namely, TP, we define this procedure as follows:

Approximation Procedure 1 (for $\boldsymbol{n_X^{**}}$, $\boldsymbol{X} \in \{\boldsymbol{e}, \boldsymbol{TP}\}$)

- Select a small number of initial measurements, n_{ini}, and evaluate $\widehat{X}(n_{ini})$.

- Keeping in mind that $\widehat{X}(n_{ini})$ may be larger than $\widehat{X}(n_X^{**}(\gamma,\delta))$, select a small safety factor $\epsilon > 0$ and calculate $\widehat{X}_{ini}^{\epsilon} = (1-\epsilon)\widehat{X}(n_{ini})$.

- Use $\widehat{X}_{ini}^{\epsilon}$ instead of $\widehat{X}$ in the right-hand side of the expression for $n_X^{**}(\gamma,\delta)$ (e.g. (3.8) in the case of $X = e$) to evaluate an approximation of $n_X^{**}(\gamma,\delta)$, denoted as $n_X^{**}(\gamma,\delta;\widehat{X}_{ini}^{\epsilon})$.

- Finally, denote the total number of measurements, thus obtained, as $\breve{n}_X^{**}(\gamma,\delta;\widehat{X}_{ini}^{\epsilon}) = \max\{n_X^{**}(\gamma,\delta;\widehat{X}_{ini}^{\epsilon}), n_{ini}\}$.

For the case of $\widehat{e}$, the effectiveness of this procedure has been investigated using Monte Carlo simulations. Specifically, to verify if $\breve{n}_e^{**}(\gamma,\delta;\widehat{e}_{ini}^{\epsilon})$ indeed results in (γ,δ)-precise estimate of e, the set of (γ,δ) pairs has been selected as

$$(\gamma,\delta) \in \{(0.1, 0.9), (0.1, 0.95), (0.05, 0.9), (0.05, 0.95)\}.$$

The parameters of the machines have been selected equiprobably from the sets

$$T_{down} \in [3, 5] \text{ and } e \in [0.6, 0.95],$$

and 25 machines have been created. As an example, five of them are shown in Table 3.3. After numerical experimentations, the values of n_{ini} and ϵ have been chosen as 30 and 0.05, respectively. Then, the performance of each machine has been simulated to obtain 10,000 realizations of $\widehat{e}_{ini}^{\epsilon}$ and corresponding realizations of $\breve{n}_e^{**}(\gamma,\delta;\widehat{e}_{ini}^{\epsilon})$ for each pair of (γ,δ). Based on the simulation results, the frequency of the event that the required (γ, δ) is observed has been calculated according to

$$\widehat{\delta} = \frac{\text{number of times } \frac{|e - \widehat{e}(\breve{n}_e^{**}(\gamma,\delta;\widehat{e}_{ini}^{\epsilon}))|}{e} \leqslant \gamma}{10000}. \tag{3.9}$$

Table 3.3. Parameters of randomly created machines.

Parameters \ Machine	m_1	m_2	m_3	m_4	m_5
T_{up}	17.99	7.73	18.89	34.37	6.69
T_{down}	3.90	3.10	3.48	3.82	3.74
e	0.82	0.71	0.84	0.90	0.64

The results obtained for the five machines of Table 3.3 are shown in Table 3.4. As one can see, all pairs of (γ, δ) have been satisfied.

Table 3.4. Values of $\widehat{\delta}$ as a function of (γ, δ).

(γ, δ) \ Machine	m_1	m_2	m_3	m_4	m_5
(0.1, 0.9)	0.9894	0.9164	0.9958	0.9980	0.9082
(0.1, 0.95)	0.9935	0.9568	0.9976	0.9995	0.9511
(0.05, 0.9)	0.9426	0.9211	0.9521	0.9889	0.9159
(0.05, 0.95)	0.9771	0.9617	0.9797	0.9922	0.9580

The values of $\breve{n}_e^{**}(\gamma, \delta; \widehat{e}_{ini}^{\epsilon})$, calculated using Approximation Procedure 1, were compared with $n_e^{**}(\gamma, \delta)$, calculated using (3.8) with $\widehat{e} = e$. The results are shown in Table 3.5 for the five machines in Table 3.3. As one can see, the difference is in the range of 20% to 60%, except for the cases where $n_e^{**}(\gamma, \delta)$ is relatively small.

The results similar to those reported in Tables 3.4 and 3.5 have been obtained for the other 20 machines analyzed (see Appendix B).

Table 3.5. Values of $n_e^{**}(\gamma,\delta)$ and $\breve{n}_e^{**}(\gamma,\delta;\widehat{e}_{ini}^{\epsilon})$ as functions of (γ,δ).

Machine / $(\boldsymbol{\gamma},\boldsymbol{\delta})$	$\boldsymbol{m_1}$		$\boldsymbol{m_2}$		$\boldsymbol{m_3}$		$\boldsymbol{m_4}$		$\boldsymbol{m_5}$	
	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$
(0.1, 0.9)	18	33	45	59	14	31	6	30	70	86
(0.1, 0.95)	25	41	63	83	19	35	8	30	99	122
(0.05, 0.9)	69	110	177	233	53	90	22	49	279	340
(0.05, 0.95)	98	156	252	332	75	127	31	69	396	485

3.4 Induced Precision and Inverse Problem for System Throughput

3.4.1 Solution of Induced Precision Problem for $\boldsymbol{TP}$

As indicated in Section 3.2, serial production lines with M exponential machines can be reduced to a single machine using the recursive aggregation procedure of *Li and Meerkov* (2009). When $T_{up,i}$ and $T_{down,i}$, $i = 1, \ldots, M$, are known precisely, the throughput of this aggregated machine is deterministic and denoted as TP. When only (α, β)-precise estimates $\widehat{T}_{up,i}$ and $\widehat{T}_{down,i}$ are available, TP is also random, and its estimate, $\widehat{TP}$, is quantified by $(\alpha_{TP}, \beta_{TP})$ as indicated in (3.2) with $X = TP$. In this section, we characterize $(\alpha_{TP}^{ind}, \beta_{TP}^{ind})$ as functions of $(\alpha,\ \beta)$ and provide a solution of the inverse problem.

Recall that in the case of a single machine, α_e^{ind} has been characterized by (3.3). To extend this formula to aggregated machines, the notion of *production system efficiency*, e_{TP}, must be introduced. This can be accomplished by defining e_{TP} as follows:

$$e_{TP} = \frac{TP}{c_M}, \tag{3.10}$$

where c_M is the capacity of the last machine in the system. Then its estimate can be evaluated as

$$\widehat{e}_{TP}(n_T^*(\alpha,\beta)) = \frac{\widehat{TP}(n_T^*(\alpha,\beta))}{c_M}. \tag{3.11}$$

This expression allows us to extend formula (3.3) to α_{TP}^{ind} as follows:

$$\alpha_{TP}^{ind} = 2\alpha(1 - \widehat{e}_{TP}(n_T^*(\alpha,\beta))). \tag{3.12}$$

As far as β_{TP}^{ind} is concerned, recall that for a single machine, β_e^{ind} is given by (3.6). This result also can be extended to the aggregated machine (in the form of a lower bound):

Proposition 3.5. *The Gaussian approximation of* β_{TP}^{ind} *is given by*

$$\beta_{TP}^{ind} > \operatorname{erf}(\alpha\sqrt{n_T^*(\alpha,\beta)}), \tag{3.13}$$

where $n_T^*(\alpha,\beta)$ *is defined by* (2.9).

Justification. See Appendix B. □

It follows from (3.12) and (3.13) that if $\widehat{e}_{TP} > 0.5$, the precision of $\widehat{TP}(n_T^*(\alpha,\beta))$ is higher than that of $\widehat{e}$ and underlining $\widehat{T}_{up,i}$ and $\widehat{T}_{down,i}$.

3.4.2 Solution of Inverse Problem for TP

As for the inverse problem, it also remains similar to that of Section 3.3, but in the sense of an upper bound:

Proposition 3.6. *For a given (γ, δ), the upper bound of the critical number of up- and downtime measurements to ensure (γ, δ)-precise estimate of TP is given by*

$$n_{TP}^{**}\left(\gamma, \delta; \frac{\widehat{TP}}{c_M}\right) < \left\lceil \left(\frac{2\left(1 - \frac{\widehat{TP}}{c_M}\right) \operatorname{erf}^{-1}(\delta)}{\gamma} \right)^2 \right\rceil, \tag{3.14}$$

where $\widehat{TP} = \widehat{TP}(n_{TP}^{**}(\gamma, \delta))$.

Justification. See Appendix B. □

Note that due to the strict inequality in (3.14), for $\widehat{e} = \frac{\widehat{TP}}{c_M}$ the number of machine up- and downtime measurements required to obtain (γ, δ)-precise estimate of TP is smaller than that for (γ, δ)-precise estimate of e (see (3.8)).

Observation 3.2. Comparing $n_T^*(\gamma, \delta)$, defined by (2.9), with $n_{TP}^{**}(\gamma, \delta)$, defined by (3.14), we conclude that for $\frac{\widehat{TP}}{c_M} > 0.5$ performance metric TP is variability contracting.

Since the right-hand side of (3.14) depends on $\widehat{TP}(n_{TP}^{**}(\gamma, \delta))$, we use the Approximation Procedure 1 to evaluate $\widehat{TP}_{ini}^{\epsilon}$ and use it subsequently in (22) to evaluate $\breve{n}_{TP}^{**}(\gamma, \delta; \widehat{TP}_{ini}^{\epsilon})$. In addition to the simulation parameters introduced in Subsection 4.2, we consider 5-machine asynchronous serial lines with buffer and machine capacities selected randomly and equiprobably from the following sets

$$N_i = \lceil r_i \max\{c_i T_{down,i}, c_{i+1} T_{down,i+1}\}\rceil,\ r_i \in [1,3],\ i=1,\ldots,4;\ c_i \in [1,2],\ i=1,\ldots,5.$$

To verify the validity of this approach, five asynchronous serial lines, denoted as s_1, ..., s_5, have been formed and used in the simulation procedure described in Subsection 4.2. As a result, we obtain that $\widehat{\delta}$ (defined by (3.9) with $\widehat{TP}$ substituted for $\widehat{e}$) is always larger than δ. The resulting $\breve{n}^{**}_{TP}(\gamma, \delta; \widehat{TP}^{\epsilon}_{ini})$, calculated using Approximation Procedure 1, is compared with $n^{**}_{TP}(\gamma, \delta)$, calculated using (3.14) with $\widehat{TP} = TP$, in Table 3.6. As one can see, the difference is in the range of 8% to 35%.

Table 3.6. Values of $n^{**}_{TP}(\gamma, \delta)$ and $\breve{n}^{**}_{TP}(\gamma, \delta; \widehat{TP}^{\epsilon}_{ini})$ as functions of (γ, δ).

System / $(\boldsymbol{\gamma}, \boldsymbol{\delta})$	s_1		s_2		s_3		s_4		s_5	
	n^{**}_{TP}	$\breve{n}^{**}_{TP}$	n^{**}_{TP}	$\breve{n}^{**}_{TP}$	n^{**}_{TP}	$\breve{n}^{**}_{TP}$	n^{**}_{TP}	$\breve{n}^{**}_{TP}$	n^{**}_{TP}	$\breve{n}^{**}_{TP}$
(0.1, 0.9)	198	214	175	196	44	59	172	192	137	155
(0.1, 0.95)	281	303	249	278	62	83	244	272	194	220
(0.05, 0.9)	790	854	699	780	173	232	688	766	547	617
(0.05, 0.95)	1121	1210	993	1106	246	330	976	1085	776	876

3.5 Induced Precision and Inverse Problem for Production Lead Time

3.5.1 Approach

Although production lead time, LT, can be evaluated using Little's formula, $LT = \frac{WIP}{TP}$, due to complexity of the analytical expression for WIP in serial lines with exponential machines and finite buffers (see *Li and Meerkov* (2009), Chapter 11), analytical solutions of the induced precision and inverse problems for LT are all but impossible to derive. Therefore, we consider here simpler systems, namely, serial production lines with M identical exponential machines and infinite buffers. The lead time in such systems has been investigated in *Meerkov and Yan* (2016) and the following approximate expression has been obtained:

$$lt = 1 + \frac{2T_{down}}{\tau}\left(\frac{1-e}{1-\rho}\right). \tag{3.15}$$

In this expression, $lt = \lim_{M\to\infty} \frac{LT}{M}$ is the *relative lead time*; e and τ are, as usual, the machine efficiency and cycle time, respectively; and $\rho = \frac{e_0}{e}$ is the *relative raw material release rate*, where e_0 is the probability of releasing a part in the system per cycle time. To ensure that $WIP < \infty$, it is assumed that $e_0 < e$. It has been shown in *Meerkov and Yan* (2016) that (3.15) provides a high accuracy approximation of lead time in systems with $M \geqslant 5$. In this section, analytical solutions of the induced precision and inverse problems are obtained using (3.15).

3.5.2 Solution of Induced Precision Problem for *LT*

When T_{up} and T_{down} are not available, their (α, β)-precise estimates $\widehat{T}_{up}(n_T^*(\alpha,\beta))$ and $\widehat{T}_{down}(n_T^*(\alpha,\beta))$ can be used instead, leading to a random variable $\widehat{lt}$ defined by:

$$\widehat{lt} = 1 + \frac{2\widehat{T}_{down}}{\tau}\left(\frac{1-\widehat{e}}{1-\widehat{\rho}}\right), \quad \widehat{e} = \frac{\widehat{T}_{up}}{\widehat{T}_{up}+\widehat{T}_{down}}, \quad \widehat{\rho} = \frac{e_0}{\widehat{e}}. \tag{3.16}$$

Similar to the previous sections, the $(\alpha_{lt}, \beta_{lt})$-precise estimate of lt is defined by the following expression:

$$P\left(\frac{|lt-\widehat{lt}(n_T^*(\alpha,\beta))|}{lt}\leqslant\alpha_{lt}\right)\geqslant\beta_{lt}. \tag{3.17}$$

Proposition 3.7. *Given* $\frac{|\widehat{T}_{up}(n_T^*(\alpha,\beta))-Tup|}{T_{up}}\leqslant\alpha$ *and* $\frac{|\widehat{T}_{down}(n_T^*(\alpha,\beta))-T_{down}|}{T_{down}}\leqslant\alpha$*, the smallest induced* α_{lt}*, which satisfies* $\frac{|\widehat{lt}(n_T^*(\alpha,\beta))-lt|}{lt}\leqslant\alpha_{lt}$ *with accuracy* $O(\alpha^2)$ *is given by*

$$\alpha_{lt}^{ind}=\frac{\widehat{lt}-1}{\widehat{lt}}\left(1+\frac{2(\widehat{\rho}-2\widehat{\rho}\widehat{e}+\widehat{e})}{1-\widehat{\rho}}\right)\alpha, \tag{3.18}$$

where $\widehat{lt}$ *is calculated based on* (3.16).

Justification. See Appendix B. □

The relationship between α_{lt}^{ind} and α can be quantified as follows:

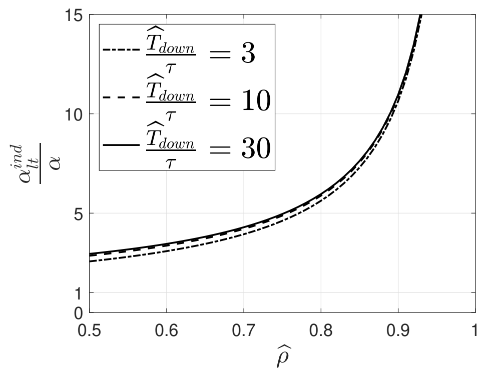
Proposition 3.8. *The inequality* $\alpha_{lt}^{ind}>\alpha$ *takes place if and only if*

$$\frac{1-\sqrt{1-\frac{\tau}{\widehat{T}_{down}}}}{2}<\widehat{e}<\frac{1+\sqrt{1-\frac{\tau}{\widehat{T}_{down}}}}{2}. \tag{3.19}$$

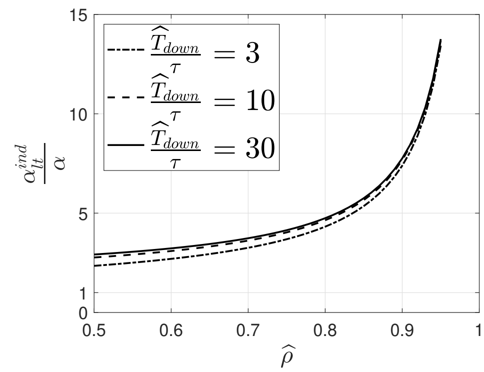
Justification. See Appendix B. □

The behavior of $\frac{\alpha_{lt}^{ind}}{\alpha}$ as a function of $\widehat{\rho}$ for various values of machine parameters is illustrated in Figure 3.2. As one can see:

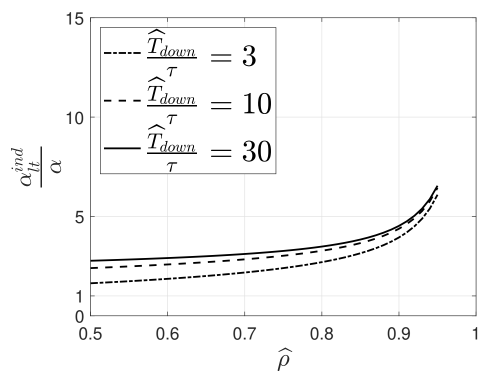
- Increasing $\widehat{\rho}$ leads to lower induced accuracy of $\widehat{lt}$;
- For $\widehat{e}>0.5$, increasing $\widehat{e}$ leads to higher induced accuracy of $\widehat{lt}$;
- The induced accuracy of $\widehat{lt}$ is relatively insensitive to $\frac{\widehat{T}_{down}}{\tau}$.



(a) $\widehat{e} = 0.5$



(b) $\widehat{e} = 0.7$



(c) $\widehat{e} = 0.9$

Figure 3.2. The ratio $\frac{\alpha_{lt}^{ind}}{\alpha}$ as a function of $\widehat{e}$, $\widehat{\rho}$ and $\frac{\hat{T}_{down}}{\tau} \in \{3, 10, 30\}$

Proposition 3.9. *The Gaussian approximation of β_{lt}^{ind} is given by*

$$\beta_{lt}^{ind} = \operatorname{erf}\left(\alpha A \sqrt{\frac{n_T^*(\alpha,\beta)}{2}}\right), \tag{3.20}$$

where

$$A = \sqrt{\frac{(1+\widehat{\rho}+2\widehat{e}-4\widehat{\rho e})^2}{(\widehat{\rho}+\widehat{e}-2\widehat{\rho e})^2+(1-2\widehat{\rho e}+\widehat{e})^2}}. \tag{3.21}$$

Justification. See Appendix B. □

Since the numerator and denominator of A can be represented as $(a+b)^2$ and a^2+b^2, respectively, with $a>0$ and $b>0$, we conclude that $A>1$. Therefore, it follows from (3.20) that

$$\beta_{lt}^{ind} > \beta. \tag{3.22}$$

3.5.3 Solution of Inverse Problem for *LT*

Proposition 3.10. *For a given (γ,δ), the critical number of up- and downtime measurements to ensure (γ,δ)-precise estimate of LT is given by*

$$n_{LT}^{**}(\gamma,\delta) = \left\lceil 2R\left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)^2\left(\frac{\operatorname{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil, \tag{3.23}$$

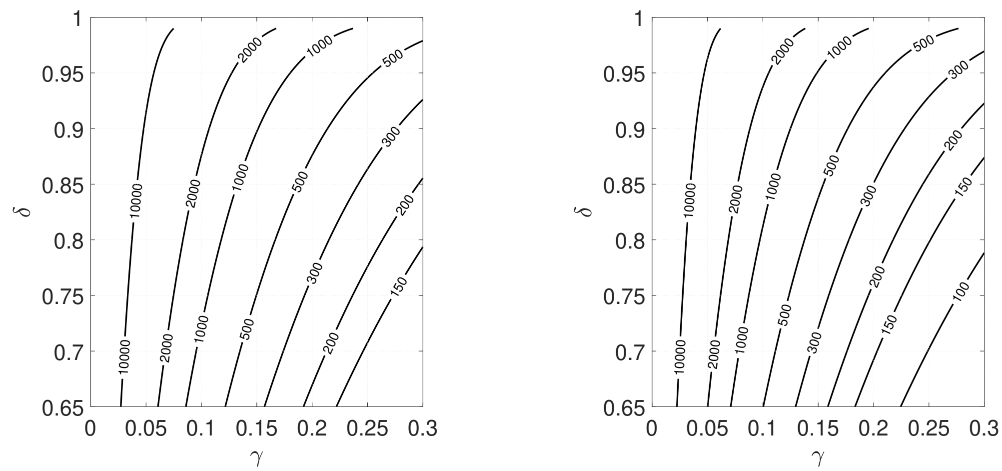
where

$$R = \frac{(\widehat{e}+\widehat{\rho}-2\widehat{\rho e})^2+(1+\widehat{e}-2\widehat{\rho e})^2}{(1-\widehat{\rho})^2}, \tag{3.24}$$

and $\widehat{lt}(n_{LT}^{**}(\gamma,\delta))$ is defined by (3.16).

Justification. See Appendix B. □

Since for $a > 0$ and $b > 0$, the numerator and denominator of R can be represented as $a^2 + b^2$ and $(a - b)^2$, respectively, $R > 1$. Therefore, for $\widehat{lt}$ sufficiently large, $n_{LT}^{**}(\gamma, \delta) > n_T^*(\gamma, \delta)$. This is illustrated by contour plots of Figure 3.3 for practical values of $\widehat{e}$, $\widehat{\rho}$, and $\frac{\widehat{T}_{down}}{\tau}$, where $n_{LT}^{**}(\gamma, \delta)$ is substantially larger than $n_T^*(\gamma, \delta)$.



(a) $\widehat{e} = 0.8$ (b) $\widehat{e} = 0.9$

Figure 3.3. Contour plots of $n_{LT}^{**}(\gamma, \delta)$ for $\widehat{e} \in \{0.8, 0.9\}$, $\widehat{\rho} = 0.8$ and $\frac{\widehat{T}_{down}}{\tau} = 10$

A question arises: Which value of $\widehat{lt}$ in (3.23) can be viewed as sufficiently large? Comparing (2.9) and (3.23), we conclude that $\widehat{lt}$ is sufficient large if

$$R \left(\frac{\widehat{lt} - 1}{\widehat{lt}} \right)^2 > 1. \tag{3.25}$$

Substituting (3.16) into this inequality, leads to the following condition:

$$\frac{\tau}{\widehat{T}_{down}} < 2 \left(\sqrt{R} - 1 \right) \frac{1 - \widehat{e}}{1 - \widehat{\rho}} \,. \tag{3.26}$$

Since the right-hand side of this inequality is monotonically increasing with respect to $\widehat{\rho}$, for every $\widehat{e}$ there exist $\widehat{\rho}^*(\widehat{e})$ such that (3.26) is satisfied. Thus, for any $\widehat{\rho} > \widehat{\rho}^*(\widehat{e})$, $\widehat{lt}$ is sufficiently large, leading to the following:

Observation 3.3. For raw material release rate sufficiently large, performance metric LT is variability expanding.

To obtain an approximation of $\widehat{lt}(n_{LT}^{**}(\gamma,\delta))$ to be used in the right-hand side of (3.23), note that in the case of variability expanding performance metrics, $n_X^{**}(\gamma,\delta) > n_T^*(\gamma,\delta)$. Therefore, n_{ini} can be selected as $n_T^*(\gamma,\delta)$, leading to the following:

Approximation Procedure 2 (for $\boldsymbol{n_X^{**}}$, $\boldsymbol{X \in \{LT, WIP\}}$)

- Select $n_{ini} = n_T^*(\gamma,\delta)$ and evaluate $\widehat{e}(n_{ini})$, $\widehat{\rho}(n_{ini})$.
- Use $\widehat{e}(n_{ini})$ and $\widehat{\rho}(n_{ini})$ instead of $\widehat{e}(n_X^{**}(\gamma,\delta))$ and $\widehat{\rho}(n_X^{**}(\gamma,\delta))$ to evaluate an approximation of $n_X^{**}(\gamma,\delta)$, denoted as $n_X^{**}(\gamma,\delta;\widehat{e}(n_{ini}),\widehat{\rho}(n_{ini}))$.
- Select a safety factor ϵ and calculate $\breve{n}_X^{**}(\gamma,\delta;\widehat{e}(n_{ini}),\widehat{\rho}(n_{ini}),\epsilon) = (1+\epsilon)n_X^{**}(\gamma,\delta;\widehat{e}(n_{ini}),\widehat{\rho}(n_{ini}))$.

To verify the validity of this approach, we use the Monte Carlo simulations with the parameters indicated in Subsection 4.2, along with $\rho \in [0.5, 0.8]$, and for five serial lines. The results are shown in Table 3.7. As one can see, the difference is quite small, in the range of 5% to 6%.

Table 3.7. Values of $n^{**}_{LT}(\gamma,\delta)$ and $\breve{n}^{**}_{LT}(\gamma,\delta;\widehat{e}(n_{ini}),\widehat{\rho}(n_{ini}),\epsilon)$ as functions of (γ,δ).

System / $(\boldsymbol{\gamma},\boldsymbol{\delta})$	s_1		s_2		s_3		s_4		s_5	
	n^{**}_{LT}	$\breve{n}^{**}_{LT}$	n^{**}_{LT}	$\breve{n}^{**}_{LT}$	n^{**}_{LT}	$\breve{n}^{**}_{LT}$	n^{**}_{LT}	$\breve{n}^{**}_{LT}$	n^{**}_{LT}	$\breve{n}^{**}_{LT}$
(0.1, 0.9)	1506	1595	1318	1389	1696	1792	584	612	520	546
(0.1, 0.95)	2138	2260	1871	1970	2408	2544	829	870	738	775
(0.05, 0.9)	6022	6337	5271	5544	6782	7133	2334	2447	2079	2183
(0.05, 0.95)	8550	8994	7484	7860	9629	10125	3313	3479	2951	3098

3.6 Induced Precision and Inverse Problem for Work-in-Process

3.6.1 Approach

Since, as it is mentioned in Subsection 3.5.1, the complexity of analytical expression for WIP prevents closed-form solutions of the induced precision and inverse problems for systems with finite buffers, we use in this section the same approach as in Section 3.5, i.e., address these problems in the framework of serial lines with M identical exponential machines and infinite buffers. For such systems, *Meerkov and Yan* (2016) provide a closed formula for work-in-process in each buffer,

$$WIP_i = \frac{2e_0 e T_{down}}{\tau}\left(\frac{1-e}{e-e_0}\right), \tag{3.27}$$

and, therefore, the total WIP in the system is

$$WIP = (M-1)\frac{2e_0 e T_{down}}{\tau}\left(\frac{1-e}{e-e_0}\right). \tag{3.28}$$

This expression is used to derive analytical solutions of the problems at hand.

3.6.2 Solution of Induced Precision Problem for $\boldsymbol{WIP}$

Proposition 3.11. *Given* $\frac{|\hat{T}_{up}(n_T^*(\alpha,\beta))-Tup|}{T_{up}} \leqslant \alpha$ *and* $\frac{|\hat{T}_{down}(n_T^*(\alpha,\beta))-T_{down}|}{T_{down}} \leqslant \alpha$, *the smallest induced* α_{WIP}, *which satisfies* $\frac{|\widehat{WIP}(n_T^*(\alpha,\beta))-WIP|}{WIP} \leqslant \alpha_{WIP}$ *with accuracy* $O(\alpha^2)$ *is given by*

$$\alpha_{WIP}^{ind} = \frac{1 + 2\widehat{e} + \widehat{\rho} - 4\widehat{e}\widehat{\rho}}{1 - \widehat{\rho}}\alpha. \tag{3.29}$$

Justification. See Appendix B. □

Subtracting denominator from numerator in (3.29), we obtain $\widehat{e} + \widehat{\rho} - 2\widehat{e}\widehat{\rho} > 0$, implying that $\alpha_{WIP}^{ind} > \alpha$.

Proposition 3.12. *The Gaussian approximation of* β_{WIP}^{ind} *is given by*

$$\beta_{WIP}^{ind} = \operatorname{erf}\left(\alpha A\sqrt{\frac{n_T^*(\alpha,\beta)}{2}}\right), \tag{3.30}$$

where A *is given by* (3.21).

Justification. See Appendix B. □

Thus, similar to the analysis of expression (3.20), we conclude that $\beta_{WIP}^{ind} > \beta$.

3.6.3 Solution of Inverse Problem for $\boldsymbol{WIP}$

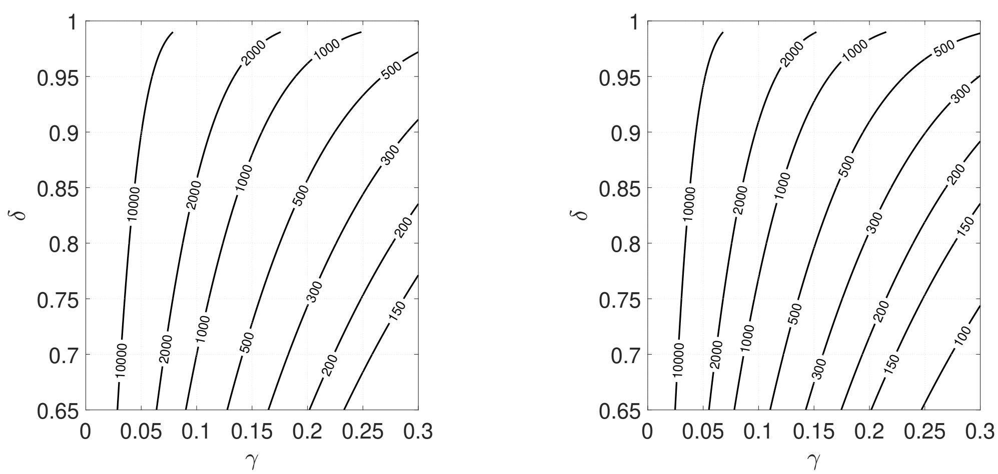
Proposition 3.13. *For a given (γ, δ), the critical number of machine up- and down-time measurements to ensure (γ, δ)-precise estimate of WIP is given by*

$$n_{WIP}^{**}(\gamma, \delta) = \left\lceil 2R \left(\frac{\text{erf}^{-1}(\delta)}{\gamma} \right)^2 \right\rceil, \tag{3.31}$$

where R is given by (3.24).

Justification. See Appendix B. □

Clearly, since $R > 1$, $n_{WIP}^{**}(\gamma, \delta)$ is larger than $n_T^{*}(\gamma, \delta)$. This is illustrated by contour plots of Figure 3.4.



(a) $\widehat{e} = 0.8$, $\widehat{\rho} = 0.8$ (b) $\widehat{e} = 0.9$, $\widehat{\rho} = 0.8$

Figure 3.4. Contour plots of $n_{WIP}^{**}(\gamma, \delta)$

Observation 3.4. The performance metric WIP is variability expanding.

Using Approximation Procedure 2, we evaluated $\breve{n}^{**}_{WIP}(\gamma, \delta; \widehat{e}(n_{ini}), \widehat{\rho}(n_{ini}), \epsilon)$ and compared it with $n^{**}_{WIP}(\gamma, \delta)$, calculated using (3.31) with $\widehat{e} = e$ and $\widehat{\rho} = \rho$. The results are shown in Table 3.8, where the difference is in the range of 5% to 7%.

Table 3.8. Values of $n^{**}_{WIP}(\gamma, \delta)$ and $\breve{n}^{**}_{WIP}(\gamma, \delta; \widehat{e}(n_{ini}), \widehat{\rho}(n_{ini}), \epsilon)$ as functions of (γ, δ).

System / $(\boldsymbol{\gamma}, \boldsymbol{\delta})$	s_1		s_2		s_3		s_4		s_5	
	n^{**}_{WIP}	$\breve{n}^{**}_{WIP}$	n^{**}_{WIP}	$\breve{n}^{**}_{WIP}$	n^{**}_{WIP}	$\breve{n}^{**}_{WIP}$	n^{**}_{WIP}	$\breve{n}^{**}_{WIP}$	n^{**}_{WIP}	$\breve{n}^{**}_{WIP}$
(0.1, 0.9)	1647	1735	2425	2582	1622	1706	1742	1838	1542	1623
(0.1, 0.95)	2338	2460	3443	3654	2303	2422	2474	2607	2189	2303
(0.05, 0.9)	6586	6920	9699	10220	6488	6818	6968	7326	6167	6479
(0.05, 0.95)	9351	9820	13771	14492	9212	9676	9893	10396	8756	9198

3.7 Summary of Formulas and Numerical Illustration

This section summarizes the formulas derived in Sections 4-8 for the critical number of measurements and provides a comparative illustration of these numbers for all performance metrics addressed.

Table 3.9 presents the formulas. The first five rows represent the performance metrics based on machines' up- and downtime measurements, and the last two – based on parts quality measurements.

Table 3.10 shows numerical values of the critical numbers, calculated using the formulas of Table 3.9, for both independent variables, T_{up}, T_{down}, q, and for functions of these variables, $X \in \{e, TP, LT, WIP\}$. As one can see, for variability contracting performance metrics, i.e., e and TP, n^{**}_X is orders of magnitude smaller than n^*_T. Conversely, for variability expanding performance metrics, i.e., LT, and WIP, n^{**}_X

is much larger than n_T^*.

Since the right-hand side of the formulas in Table 3.9 (except for the first one) depends on the performance metrics estimates, for practical utilization of these formulas, Approximation Procedures 1 and 2 (see Subsections 4.2 and 7.3) must be used. The resulting critical numbers, $\breve{n}_X^{**}$ are shown in Table 3.11. Although these numbers are larger than those of Table 3.10, for variability contracting performance metrics they are at least an order of magnitude smaller than n_T^*. Thus, the formulas of Table 3.9, combined with Approximation Procedures 1 and 2, offer a practical method for calculating the number of independent variables measurements to ensure a desired precision of all performance metrics considered in this chapter.

Table 3.9. Formulas for critical number of measurements for various performance metrics.

Metrics	**Critical Number of Measurements**
$\boldsymbol{T_{up}}$ **and** $\boldsymbol{T_{down}}$	$n_T^*(\gamma,\delta) = \left\lceil 2\left(\frac{\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2 \right\rceil$
$\boldsymbol{e}$	$n_e^{**}(\gamma,\delta) = \left\lceil \left(\frac{2(1-\hat{e})\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2 \right\rceil$
$\boldsymbol{TP}$	$n_{TP}^{**}(\gamma,\delta) < \left\lceil \left(\frac{2(1-\hat{e}_{TP})\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2 \right\rceil,\ \hat{e}_{TP} = \frac{\widehat{TP}}{c_M}$
$\boldsymbol{LT}$	$n_{LT}^{**}(\gamma,\delta) = \left\lceil 2R\left(\frac{\hat{lt}-1}{\hat{lt}}\right)^2 \left(\frac{\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2 \right\rceil,\ R = \frac{(\hat{e}+\hat{\rho}-2\widehat{\rho e})^2+(1+\hat{e}-2\widehat{\rho e})^2}{(1-\hat{\rho})^2}$
$\boldsymbol{WIP}$	$n_{WIP}^{**}(\gamma,\delta) = \left\lceil 2R\left(\frac{\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2 \right\rceil,\ R = \frac{(\hat{e}+\hat{\rho}-2\widehat{\rho e})^2+(1+\hat{e}-2\widehat{\rho e})^2}{(1-\hat{\rho})^2}$

Table 3.10. Critical number of measurements calculated using formulas of Table 3.9.

Critical Numbers \ (γ, δ)	$(0.1, 0.9)$	$(0.1, 0.95)$	$(0.05, 0.9)$	$(0.05, 0.95)$
$n_T^*(\gamma, \delta)$	271	385	1083	1537
$n_e^{**}(\gamma, \delta)$, $\hat{e} = 0.9$	6	8	22	31
$n_{TP}^{**}(\gamma, \delta)$, $\widehat{\frac{TP}{c_M}} = 0.9$, $M = 5$	5	6	9	12
$n_{LT}^{**}(\gamma, \delta), \hat{e} = 0.9, \hat{\rho} = 0.8, \frac{\hat{T}_{down}}{\tau} = 10$	1561	2216	6243	8864
$n_{WIP}^{**}(\gamma, \delta), \hat{e} = 0.9, \hat{\rho} = 0.8$	1889	2682	7754	10726

Table 3.11. Critical number of measurements calculated using Approximation Procedures 1 and 2.

Approximation of Critical Numbers \ (γ, δ)	$(0.1, 0.9)$	$(0.1, 0.95)$	$(0.05, 0.9)$	$(0.05, 0.95)$
$n_T^*(\gamma, \delta)$	271	385	1083	1537
$\breve{n}_e^{**}(\gamma, \delta; \hat{e}_{ini}^{\epsilon})$, $\hat{e} = 0.9$	30	30	49	69
$\breve{n}_{TP}^{**}(\gamma, \delta; \widehat{TP}_{ini}^{\epsilon})$, $\widehat{\frac{TP}{c_M}} = 0.9$, $M = 5$	30	30	30	35
$\breve{n}_{LT}^{**}(\gamma, \delta; \hat{e}(n_{ini}), \hat{\rho}(n_{ini}), \epsilon)$, $\hat{e} = 0.9, \hat{\rho} = 0.8, \frac{\hat{T}_{down}}{\tau} = 10$	1643	2333	6562	9320
$\breve{n}_{WIP}^{**}(\gamma, \delta; \hat{e}(n_{ini}), \hat{\rho}(n_{ini}), \epsilon)$, $\hat{e} = 0.9, \hat{\rho} = 0.8$	1991	2825	7940	11274

3.8 Summary

This chapter provides an analytical characterization of the number of measurements required for evaluating various production systems performance metrics, which are necessary for managerial decision-making and evaluating efficacy of potential continuous improvement projects. As far as the metrics are concerned, they are machine efficiency, throughput (with and without taking into account parts quality), production lead time, and work-in-process. As far as the measurements are concerned, they are duration of machines' up- and downtime. Both quantitative and qualitative characterizations of these critical numbers are provided. The quantitative ones are represented by closed formulas for the precision of the performance metrics as functions of their arguments (i.e., measurements). The qualitative ones are formulated in terms of variability of performance metrics as compared with variability of their arguments. In this regard, it is shown that under practice-inspired conditions, the variability contracting metrics are machine efficiency and throughput, while the throughput of non-defective parts produced, lead time, and work-in-process are variability expanding.

Future work in this area includes extending the results obtained to:

- assembly systems;
- production systems with machines obeying non-exponential reliability models.

Additionally, utilization of the results obtained in practice is an important part of future work.

CHAPTER 4

The (α_q, β_q)- and $(\alpha_{TP_q}, \beta_{TP_q})$-Precise Estimate of Machine Quality Parameter q and Quality Parts Throughput TP_q

4.1 Problem Motivation

The machine quality parameter q is used in quantitative method for performance analysis, continuous improvement, and design of production systems with non-perfect quality machines. Such machines produce a non-defective part (i.e., quality part) with probability q, and defective with probability $1 - q$. To evaluate q for non-perfect machines, parts quality must be measured and then averaged to obtain the estimate $\widehat{q}$. The law of large number guarantees that a sufficiently large number of measurements leads to sufficiently precise estimate. However, collecting a large number of measurements requires a long observation time. This chapter is intended to provide guidance for selecting the smallest number of parts quality measurements necessary and sufficient for calculating reliable estimate of q. The term "reliable" is used to indicate an estimate, which has the desired accuracy with the desired probability.

In addition to q, this chapter addresses one performance metric, namely, quality parts throughput (TP_q) in production systems with non-perfect quality machines obeying the Bernoulli quality model. A method for performance analysis of systems with non-perfect quality machines and inspection stations (intended to remove defective parts or direct them for rework) is developed in *Li and Meerkov* (2009). On this basis, this chapter calculates induced $(\alpha_{TP_q}, \beta_{TP_q})$-precise estimate of TP_q and the smallest number of parts quality measurements required to ensure the desired precision of $\widehat{TP}_q$.

Along with quantitative results, this chapter provides qualitative insights into variability properties of production system quality parts throughput. To describe it, let $n_q^*(\gamma, \delta)$ denote the smallest number of up- and downtime measurements required to obtain (γ, δ)-precise estimates of q, while $n_{TP_q}^{**}(\gamma, \delta)$ denotes the smallest number of up- and downtime measurements required to obtain an equally precise estimate of TP_q.

Definition 4.1. Performance metric TP_q is:

- *variability contracting* if $n_{TP_q}^{**}(\gamma, \delta) < n_q^*(\gamma, \delta)$;
- *variability expanding* if $n_{TP_q}^{**}(\gamma, \delta) > n_q^*(\gamma, \delta)$.

4.2 Problems Formulation

Consider a non-perfect quality exponential machine producing quality parts with probability q and defectives with probability $1-q$. Assume the machine produced n parts, of which n_q were non-defective. Then, its quality parameter q can be estimated as

$$\widehat{q}(n) = \frac{n_q(n)}{n}. \tag{4.1}$$

Definition 4.2.

- The estimate $\widehat{q}(n)$ is referred to as (α_q, β_q)-*precise* if

$$P\left\{\frac{|q-\widehat{q}(n)|}{q} \leqslant \alpha_q\right\} \geqslant \beta_q. \tag{4.2}$$

- The smallest integer $n_q^*(\alpha_q, \beta_q)$, which guarantees the above relationship, is referred to as the *critical number of parts quality measurements.*

The $\boldsymbol{n_q^*(\alpha_q, \beta_q)}$ evaluation problem: Given a desired (α_q, β_q), calculate the critical number of parts quality measurements, $n_q^*(\alpha_q, \beta_q)$.

Consider now a serial production line consisting of M exponential machines with the quality parameter $q_i \leqslant 1$, $i = 1, \ldots, M$. Assume that this system is equipped with an inspection station at the end of the line. In this case, if the values of $T_{up,i}$, $T_{down,i}$, and q_i are available, the throughput of quality parts, TP_q, can be calculated as (see *Meerkov and Zhang* (2010))

$$TP_q = TP \prod_{i=1}^{M} q_i, \tag{4.3}$$

where TP is the throughput of the line calculated under the assumption that all machines are of perfect quality (i.e., produce no defective parts). When $T_{up,i}$ and $T_{down,i}$, $i = 1, \ldots, M$, are known, but only (α_q, β_q)-precise estimates of q_i, $i = 1, \ldots, M$, are available, the estimate of TP_q can be evaluated as

$$\widehat{TP}_q = TP \prod_{i=1}^{M} \widehat{q}_i. \tag{4.4}$$

Finally, when neither $T_{up,i}$ and $T_{down,i}$ nor q_i, $i = 1, \ldots, M$, are known precisely and only their estimates are available, the estimate of TP_q is

$$\widehat{\widehat{TP}}_q = \widehat{TP} \prod_{i=1}^{M} \widehat{q}_i, \tag{4.5}$$

where $\widehat{TP}$ is calculated assuming no defective parts are produced.

The induced accuracy of both $\widehat{TP}_q$ and $\widehat{\widehat{TP}}_q$ can be evaluated. Due to space limitations, only the former is addressed in this chapter. Specifically, quantifying the induced $(\alpha_{TP_q}, \beta_{TP_q})$-precision of $\widehat{TP}_q(n_q^*(\alpha_q, \beta_q))$ as

$$P\left\{\frac{|TP_q - \widehat{TP}_q(n_q^*(\alpha_q, \beta_q))|}{TP_q} \leqslant \alpha_{TP_q}\right\} \geqslant \beta_{TP_q}, \tag{4.6}$$

we introduce the following problems:

Induced precision problem for $\boldsymbol{TP_q}$: Given (α_q, β_q)-precise estimates $\widehat{q}_i(n_q^*(\alpha_q, \beta_q))$, $i = 1, \ldots, M$, calculate the induced precision $(\alpha_{TP_q}^{ind}, \beta_{TP_q}^{ind})$ of $\widehat{TP}_q$.

Inverse problem for $\boldsymbol{TP_q}$: Given a desired (γ, δ), calculate the smallest number of parts quality measurements, $n_{TP_q}^{**}(\gamma, \delta)$, required to obtain (γ, δ)-precise estimate

$\widehat{TP}_q$.

4.3 The $(\boldsymbol{\alpha_q}, \boldsymbol{\beta_q})$-precise Estimate of Machine Quality Parameter

The (α_q, β_q)-precise estimate of the machine quality parameter q is introduced in Definition 4.2 by expression (4.2) along with the critical number of parts quality measurements, $n_q^*(\alpha_q, \beta_q)$, required for (4.2) to take place. The proposition below provides an expression for this number.

Proposition 4.1. *The Gaussian approximation of the critical number of parts quality measurements to ensure* (α_q, β_q)*-precise estimate of* q *is given by*

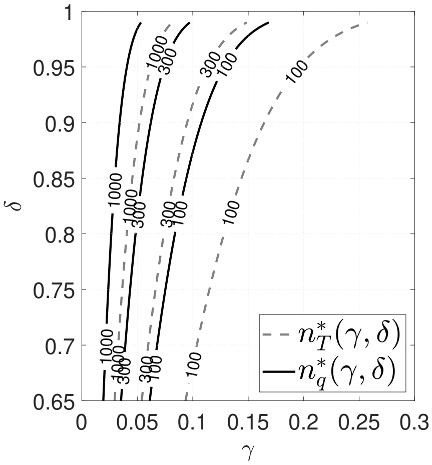
$$n_q^*(\alpha_q, \beta_q) = \left\lceil 2 \left(\frac{1-\widehat{q}}{\widehat{q}} \right) \left(\frac{\mathrm{erf}^{-1}(\beta_q)}{\alpha_q} \right)^2 \right\rceil, \tag{4.7}$$

where $\widehat{q} = \widehat{q}(n_q^*(\alpha_q, \beta_q))$.

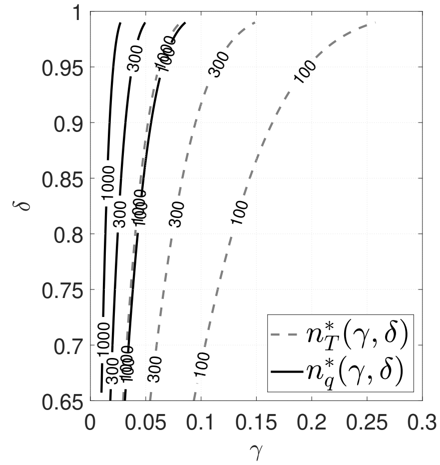
Justification. See Appendix C. □

As it follows from (2.9) and (4.7), $n_q^*(\gamma, \delta) < n_T^*(\gamma, \delta)$ if $\widehat{q} > 0.5$. This is a result of the fact that the coefficient of variation of a Bernoulli random variable (parts quality) is smaller than the coefficient of variation of exponential random variable (machines' up- and downtime) if $\widehat{q} > 0.5$. This phenomenon is illustrated in Figure 4.1, and the contour plots of $n_q^*(\gamma, \delta; \widehat{q})$ are shown in Figure 4.2.

As for all previously considered performance metrics, the dependence of the right-hand side of (4.7) on $\widehat{q}(n_q^*(\gamma, \delta))$ can be eliminated using Approximation Procedure

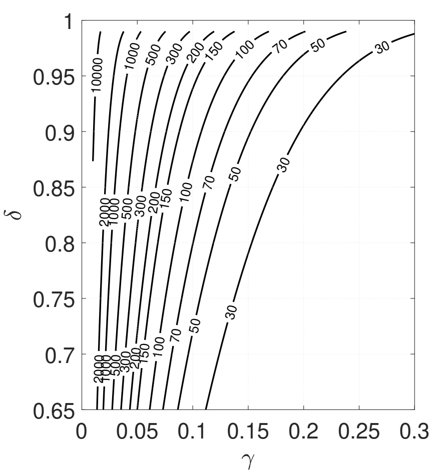


(a) $\widehat{q} = 0.7$

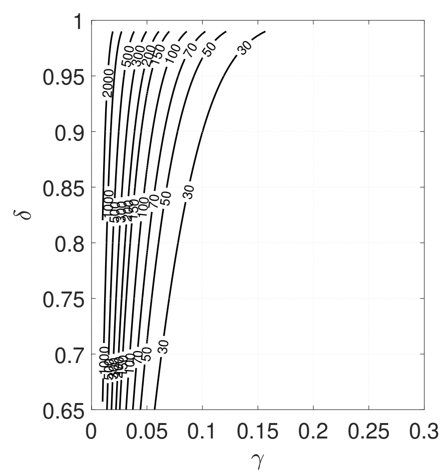


(b) $\widehat{q} = 0.9$

Figure 4.1. Comparison of $n_q^*(\gamma, \delta)$ and $n_T^*(\gamma, \delta)$



(a) $\widehat{q} = 0.7$



(b) $\widehat{q} = 0.9$

Figure 4.2. Contour plots of $n_q^*(\gamma, \delta; \widehat{q})$

1. However, as it turns out, for the case of q, this procedure must be modified by having safety factor ϵ equal to 0.1, rather than 0.05. The reason is that for $\epsilon = 0.05$, the number of measurements $\breve{n}_q^*(\gamma, \delta; \widehat{q}_{ini}^\epsilon)$ turns out to be too small to guarantee that $\widehat{\delta}$, defined in (3.9) with q substituted instead of e is smaller than δ, while for $\epsilon = 0.1$, $\widehat{\delta}$ is indeed larger than δ. The values of $\breve{n}_q^*(\gamma, \delta; \widehat{q}_{ini}^\epsilon)$, calculated according to Approximation Procedure 1 with $\epsilon = 0.1$ and q selected randomly and equiprobably from the set $[0.6, 0.95]$, are shown in Table 4.1, along with $n_q^*(\gamma, \delta)$. As one can see, due to a larger safety factor, the difference is significant, ranging from 40% to 150%.

Table 4.1. Values of $n_q^*(\gamma, \delta)$ and $\breve{n}_q^*(\gamma, \delta; \widehat{q}_{ini}^\epsilon)$ as functions of (γ, δ).

Machine / $(\boldsymbol{\gamma}, \boldsymbol{\delta})$	$\boldsymbol{m_1}$		$\boldsymbol{m_2}$		$\boldsymbol{m_3}$		$\boldsymbol{m_4}$		$\boldsymbol{m_5}$	
	$\boldsymbol{q_1 = 0.86}$		$\boldsymbol{q_2 = 0.68}$		$\boldsymbol{q_3 = 0.85}$		$\boldsymbol{q_4 = 0.80}$		$\boldsymbol{q_5 = 0.92}$	
	n_q^*	$\breve{n}_q^*$	n_q^*	$\breve{n}_q^*$	n_q^*	$\breve{n}_q^*$	n_q^*	$\breve{n}_q^*$	n_q^*	$\breve{n}_q^*$
(0.1, 0.9)	45	82	129	183	46	84	67	108	24	58
(0.1, 0.95)	64	116	184	257	66	119	95	154	34	82
(0.05, 0.9)	180	330	516	722	184	334	268	431	94	229
(0.05, 0.95)	255	469	733	1031	261	474	380	614	133	323

4.4 Induced Precision and Inverse Problem for Throughput of Non-defective Parts

4.4.1 Solution of Induced Precision Problem for $\boldsymbol{TP_q}$

As indicated in Section 4.2, the estimate of TP_q is defined by (4.4) and its $(\alpha_{TP_q}, \beta_{TP_q})$ precision by (4.6). In terms of these definitions, we obtain:

Proposition 4.2. *Given* $\frac{|q_i - \widehat{q}_i(n_q^*(\alpha_q,\beta_q))|}{q_i} \leqslant \alpha_q$, $i = 1,\ldots,M$, *the smallest* α_{TP_q}, *which satisfies* $\frac{|TP_q - \widehat{TP}_q(n_T^*(\alpha_q,\beta_q))|}{TP_q} \leqslant \alpha_{TP_q}$ *with accuracy up to* $O(\alpha_q^2)$ *is given by*

$$\alpha_{TP_q}^{ind} = k\alpha_q, \tag{4.8}$$

where k *is number of non-perfect quality machines in a serial line.*

Justification. See Appendix C. □

Thus, the induced accuracy of $\widehat{TP}_q$ decreases linearly with the number of non-perfect quality machines.

Proposition 4.3. *Assume* $\prod_{i=1}^{M} q_i > 0.5$. *Then, the Gaussian approximation of* $\beta_{TP_q}^{ind}$ *is given by*

$$\beta_{TP_q}^{ind} > \operatorname{erf}\left(\frac{k\alpha_q}{\sqrt{2}}\sqrt{n_q^*(\alpha_q,\beta_q)}\right). \tag{4.9}$$

Justification. See Appendix C. □

Thus, the induced probability of $\widehat{TP}_q$ is increasing with the number of non-perfect quality machines, however, in a nonlinear manner.

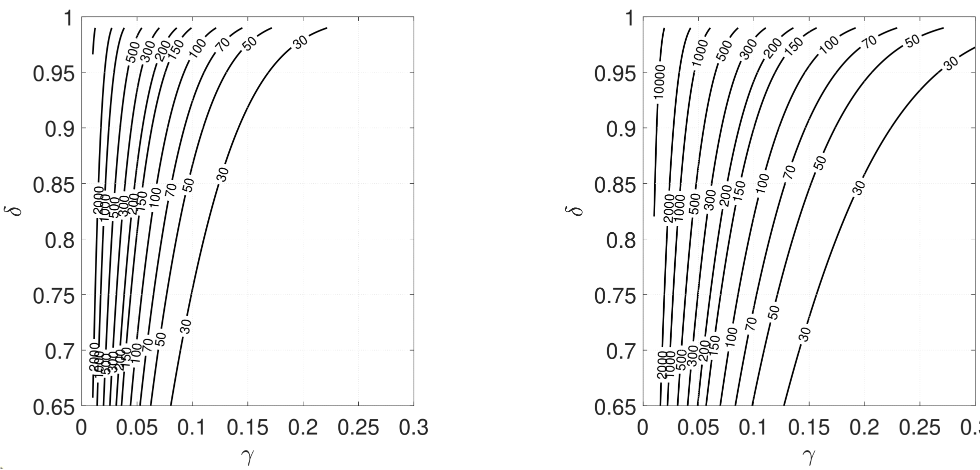
4.4.2 Solution of Inverse Problem for TP_q

Proposition 4.4. *For a given* (γ,δ), *the critical number of parts quality measurements to ensure* (γ,δ)*-precise estimate of* TP_q *is given by*

$$n_{TP_q}^{**}(\gamma,\delta) = \left\lceil 2\left(\sum_{i=1}^{k}\frac{1-\widehat{q}_i}{\widehat{q}_i}\right)\left(\frac{\operatorname{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil. \tag{4.10}$$

Justification. See Appendix C. □

As it follows from (4.7) and (4.10), $n^{**}_{TP_q}(\gamma,\delta) > n^*_q(\gamma,\delta)$ for $k \geqslant 2$. The contour plots of $n^{**}_{TP_q}(\gamma,\delta)$ are shown in Figure 4.3.



(a) $k = 2$ (b) $k = 5$

Figure 4.3. Contour plots of $n^{**}_{TP_q}(\gamma,\delta)$ for $\widehat{q}_i = 0.9$, $i = 1,\dots,k$

Observation 4.1. As it follows from the above, performance metric TP_q for serial lines with more than one non-perfect quality machines is variability expanding.

The values of $\breve{n}^{**}_{TP_q}(\gamma,\delta;\widehat{TP}^{\epsilon}_{q,ini})$ is compared with $n^{**}_{TP_q}(\gamma,\delta)$, calculated using (4.10) with $\widehat{q}_i = q_i$, in Table 4.2 for five asynchronous serial lines with $q_i \in [0.6, 0.95]$, $i = 1,\dots,5$, and $\prod_{i=1}^{5} q_i > 0.5$. As one can see, the difference is in the range of 40% to 50%.

Table 4.2. Values of $n^{**}_{TP_q}(\gamma,\delta)$ and $\breve{n}^{**}_{TP_q}(\gamma,\delta;\widehat{TP}^{\epsilon}_{q,ini})$ as functions of (γ,δ).

System / $(\boldsymbol{\gamma},\boldsymbol{\delta})$	s_1		s_2		s_3		s_4		s_5	
	$n^{**}_{TP_q}$	$\breve{n}^{**}_{TP_q}$	$n^{**}_{TP_q}$	$\breve{n}^{**}_{TP_q}$	$n^{**}_{TP_q}$	$\breve{n}^{**}_{TP_q}$	$n^{**}_{TP_q}$	$\breve{n}^{**}_{TP_q}$	$n^{**}_{TP_q}$	$\breve{n}^{**}_{TP_q}$
(0.1, 0.9)	201	292	192	281	206	297	194	285	177	265
(0.1, 0.95)	285	413	272	398	293	421	276	404	251	376
(0.05, 0.9)	802	1169	765	1118	824	1192	776	1135	705	1053
(0.05, 0.95)	1138	1651	1086	1599	1169	1690	1102	1614	1001	1500

4.5 Summary of Formulas and Numerical Illustration

This section summarizes the formulas derived in Sections 4-8 for the critical number of parts quality measurements and provides a comparative illustration of these numbers.

Table 4.3 presents the formulas. Table 4.4 shows numerical values of the critical numbers, calculated using the formulas of Table 4.3. As one can see, for variability expanding performance metrics TP_q, $n^{**}_{TP_q}$ is much larger than n^*_q.

Since the right-hand side of the formulas in Table 3.9 (except for the first one) depends on the performance metrics estimates, for practical utilization of these formulas, Approximation Procedures 1 in Chapter 3 must be used. The resulting critical numbers are shown in Table 4.5.

4.6 Summary

This chapter provides both quantitative and qualitative characterizations of the critical number of parts quality measurements. As far as the quantitative char-

Table 4.3. Formulas for critical number of measurements for various performance metrics.

Metrics	Critical Number of Measurements
$\boldsymbol{q}$	$n_q^*(\gamma,\delta) \quad = \left\lceil 2\left(\frac{1-\widehat{q}}{\widehat{q}}\right)\left(\frac{\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil$
$\boldsymbol{TP_q}$	$n_{TP_q}^{**}(\gamma,\delta) \quad = \left\lceil 2\left(\sum_{i=1}^{k}\frac{1-\widehat{q}_i}{\widehat{q}_i}\right)\left(\frac{\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil$

Table 4.4. Critical number of measurements calculated using formulas of Table 4.3.

Critical Numbers \ $(\boldsymbol{\gamma},\boldsymbol{\delta})$	$(\mathbf{0.1},\mathbf{0.9})$	$(\mathbf{0.1},\mathbf{0.95})$	$(\mathbf{0.05},\mathbf{0.9})$	$(\mathbf{0.05},\mathbf{0.95})$
$\boldsymbol{n_q^*(\gamma,\delta),\ \widehat{q}=0.9}$	31	43	121	171
$\boldsymbol{n_{TP_q}^{**}(\gamma,\delta),\widehat{q}_i=0.9, i=1,...,5, M=5}$	151	214	602	854

Table 4.5. Critical number of measurements calculated using Approximation Procedures 1 of Chapter 3.

Approximation of Critical Numbers \ $(\boldsymbol{\gamma},\boldsymbol{\delta})$	$(\mathbf{0.1},\mathbf{0.9})$	$(\mathbf{0.1},\mathbf{0.95})$	$(\mathbf{0.05},\mathbf{0.9})$	$(\mathbf{0.05},\mathbf{0.95})$
$\boldsymbol{\breve{n}_q^*(\gamma,\delta;\widehat{q}_{ini}^{\epsilon}),\ \widehat{q}=0.9}$	65	92	258	367
$\boldsymbol{\breve{n}_{TP_q}^{**}(\gamma,\delta;\widehat{TP}_{q,ini}^{\epsilon}),}$ $\boldsymbol{\widehat{q}_i=0.9, i=1,...,5, M=5}$	236	337	942	1339

acterization is concerned, this chapter calculates the critical number $n_q^*(\alpha_q, \beta_q)$ to ensure (α_q, β_q)-precise estimate of $\widehat{q}$, evaluates the induced precision of the quality parts throughput estimate $\widehat{TP}_q(n^*(\alpha_q, \beta_q))$, and presents the solution to the inverse problem, i.e., the critical number $n_{TP_q}^{**}(\gamma, \delta)$ to ensure (γ, δ)-precise estimate of $\widehat{TP}_q$. Formulas and values of both critical numbers are summarized in Table 4.3-4.5. As far as the qualitative characterization is concerned, this chapter shows that TP_q is a variability expanding performance metric, i.e, the for obtaining the same level of precision, the critical number of parts quality measurement required for the estimate of TP_q is larger than that for the estimate of q.

CHAPTER 5

Comparison of the Critical Numbers Calculated Using (α, β)-Precision Theory with those Evaluated Using Markov Inequality, Chebyshev Inequality, and by Simulations

5.1 Problem Motivation

In Chapters 2 and 4, the critical number of measurements required to obtain (α, β)-precise estimates of $\widehat{T}_{up}$, $\widehat{T}_{down}$, $\widehat{q}$ defined by Definitions 2.1 and 4.2 has been calculated using the technique developed in this dissertation, i.e., (α, β)-Precision Theory. Since $\widehat{T}_{up}$ and $\widehat{T}_{down}$ are random variables, their statistical characterization can also be obtained using the classical probabilistic inequalities, namely, Markov and Chebyshev inequalities, as well as by numerical simulations. This leads to a possibility of evaluating the critical numbers of measurements based on these tools. The purpose of this chapter is to carry out such evaluations and compare the results with those derived in Chapters 2 and 4.

For the probabilistic inequalities, we use the Markov's and Chebyshev's inequalities, which are often used in probability related calculations, see, for instance, *Boucheron et al.* (2003); *Gubner* (2006); *Resnick* (2019).

As defined in Definition 2.1, the estimates $\widehat{T}_{up}$ and $\widehat{T}_{down}$ are (α, β)-precise if they satisfy the following inequalities:

$$P\left\{\frac{|\widehat{T}_{up}(n) - T_{up}|}{T_{up}} \leqslant \alpha\right\} \geqslant \beta, \quad P\left\{\frac{|\widehat{T}_{down}(n) - T_{down}|}{T_{down}} \leqslant \alpha\right\} \geqslant \beta.$$

The critical number of measurements obtained using the (α, β)-Precision Theory has been investigated in Chapter 2 and the following expression has been obtained:

$$n_T^*(\alpha, \beta) = \left\lceil 2\left(\frac{\text{erf}^{-1}(\beta)}{\alpha}\right)^2\right\rceil. \tag{5.1}$$

Denote the critical numbers of measurements evaluated based on Markov inequality, Chebyshev inequality, and by simulations as $n_{T,M}^*(\alpha, \beta)$, $n_{T,C}^*(\alpha, \beta)$, and $n_{T,S}^*(\alpha, \beta)$, respectively.

Similarly, the estimate $\widehat{q}$ is (α_q, β_q)-precise if it satisfy

$$P\left\{\frac{|\widehat{q}(n) - q|}{q} \leqslant \alpha_q\right\} \geqslant \beta_q.$$

Using the (α, β)-Precision Theory, we obtain the critical number of parts quality measurement, given by:

$$n_q^*(\alpha_q, \beta_q) = \left\lceil 2\left(\frac{1-\widehat{q}}{\widehat{q}}\right)\left(\frac{\mathrm{erf}^{-1}(\beta_q)}{\alpha_q}\right)^2 \right\rceil.$$

Similar to the ones for T_{up} and T_{down}, we denote the critical numbers of parts quality measurements evaluated based on Markov inequality, Chebyshev inequality, and by simulations as $n_{q,M}^*(\alpha, \beta)$, $n_{q,C}^*(\alpha, \beta)$, and $n_{q,S}^*(\alpha, \beta)$, respectively.

The evaluations of those critical numbers are given in Sections 5.2-5.4.

5.2 Critical Numbers Evaluated Using Markov Inequality

Proposition 5.1. *The Gaussian approximation of critical number $n_{T,M}^*(\alpha, \beta)$ is given by*

$$n_{T,M}^*(\alpha, \beta) = \left\lceil \frac{2}{\pi\alpha^2(1-\beta)^2} \right\rceil.$$

Justification. See Appendix D. □

Proposition 5.2. *The Gaussian approximation of the critical number $n_{q,M}^*(\alpha_q, \beta_q)$ is given by*

$$n_{q,M}^*(\alpha_q, \beta_q) = \left\lceil \frac{2(1-q)}{\pi q} \frac{1}{\alpha_q^2(1-\beta_q)^2} \right\rceil.$$

Justification. See Appendix D. □

5.3 Critical Numbers Evaluated Using Chebyshev Inequality

Theorem 5.3. *The critical number* $n^*_{T,C}(\alpha, \beta)$ *is given by*

$$n^*_{T,C}(\alpha, \beta) = \left\lceil \frac{1}{\alpha^2(1-\beta)} \right\rceil.$$

Proof. See Appendix D. □

Theorem 5.4. *The critical number* $n^*_{q,C}(\alpha_q, \beta_q)$ *is given by*

$$n^*_{q,C}(\alpha_q, \beta_q) = \left\lceil \frac{1-q}{q} \frac{1}{\alpha_q^2(1-\beta_q)} \right\rceil.$$

Proof. See Appendix D. □

5.4 Critical Numbers Evaluated Using Simulation

For any number of observation n and T_{up}, we randomly generate n number of exponential random variable with mean T_{up}, and compute the mean of the realizations, denoted as $\widehat{T}_{up}(n)$. This procedure is repeated for $N = 10000$ times. The probability β can be approximated by

$$\widehat{\beta} = \frac{\text{number of times } \frac{|\hat{T}_{up}(n) - T_{up}|}{T_{up}} \leqslant \alpha}{N}.$$

Using the approximation $\widehat{\beta}$, the critical number $n^*_{T,S}(\alpha, \beta)$ can be evaluated by a line search.

For any number of observation n and q, we randomly generate n number of Bernoulli random variable with parameter q, and compute the mean of the realizations, denoted as $\widehat{q}(n)$. This procedure is repeated for $N = 10000$ times. The probability β_q can be approximated by

$$\widehat{\beta}_q = \frac{\text{number of times } \frac{|\widehat{q}(n)-q|}{q} \leqslant \alpha_q}{N}.$$

Using the approximation $\widehat{\beta}_q$, the critical number $n^*_{q,S}(\alpha_q, \beta_q)$ can be evaluated by a line search.

5.5 Numerical Illustration

For $T_{up} = 10$ and the precision pair $(\alpha, \beta) \in \{(0.1, 0.9), (0.1, 0.95), (0.05, 0.9), (0.05, 0.95)\}$, the corresponding critical numbers $n^*_T(\alpha, \beta), n^*_{T,M}(\alpha, \beta), n^*_{T,S}(\alpha, \beta)$ and $n^*_{T,C}(\alpha, \beta)$ are listed in Table 5.1.

Table 5.1. Critical numbers $n^*_{T,M}(\alpha, \beta), n^*_{T,C}(\alpha, \beta), n^*_{T,S}(\alpha, \beta)$ and $n^*_T(\alpha, \beta)$ as functions of (α, β).

Critical Number \ (α, β)	$(0.1, 0.9)$	$(0.1, 0.95)$	$(0.05, 0.9)$	$(0.05, 0.95)$
$n^*_{T,M}(\alpha, \beta)$	6366	25465	25465	101860
$n^*_{T,C}(\alpha, \beta)$	1000	2000	4000	8000
$n^*_{T,S}(\alpha, \beta)$	271	386	1084	1543
$n^*_T(\alpha, \beta)$	271	385	1083	1537

Notice that, since q is unknown, the expressions for $n^*_{q,M}(\alpha_q, \beta_q)$ and $n^*_{q,C}(\alpha_q, \beta_q)$ are not closed and $n^*_{q,S}(\alpha_q, \beta_q)$ can not be evaluated. Thus, the same approximation procedure for $n_q^*(\alpha_q, \beta_q)$ is needed for $n_{q,M}^*(\alpha_q, \beta_q), n_{q,C}^*(\alpha_q, \beta_q)$, and $n_{q,S}^*(\alpha_q, \beta_q)$. We take the same $n_{ini} = 30$ and $\epsilon = 0.1$ as we have for $\check{n}_q^*(\alpha_q, \beta_q)$.

For $q = 0.9$ and the precision pair $(\alpha_q, \beta_q) \in \{(0.1, 0.9),\ (0.1, 0.95),\ (0.05, 0.9),\ (0.05, 0.95)\}$, the critical numbers $n_q^*(\alpha_q, \beta_q), n_{q,M}^*(\alpha_q, \beta_q)$ and $n_{q,C}^*(\alpha_q, \beta_q)$ and the corresponding approximated values $\check{n}_{q,M}(\alpha_q, \beta_q), \check{n}_{q,C}(\alpha_q, \beta_q), \check{n}_{q,S}(\alpha_q, \beta_q), \check{n}_q(\alpha_q, \beta_q)$ are listed in the following table. As we can see, the approximated critical numbers are about twice the actual value.

The time complexity of the simulation method is $\mathcal{O}\left(Nn\log(n)\right)$, where n is the upper bound of the critical number used in the simulations (e.g., $n_{T,C}^*$), while the time complexity of the (α, β)-Precision Theory is only $\mathcal{O}(1)$. For comparison, the computation time for obtaining $n_{T,S}^*$, n_T^*, $n_{q,S}^*$, and n_q^* of Tables 5.1 and 5.2 are reported in Table 5.3. As we can see, the simulation approach can take up to several seconds to evaluate a critical number, which is much slower than computing the value using closed-form formula. For reference, the simulations are conducted on a Windows computer with 2-core Intel Core i7-6500U CPU @ 2.50GHz, 8GB 1866MHz DDR3 RAM using MATLAB 2020b. Other disadvantages of the simulation method include it cannot provide analytical formulas of the critical numbers as a functions of either α or β, and simulations have to repeat anew for every value of machine parameters.

5.6 Summary

For the evaluation of both the critical number of up- and downtime measurements and critical number of parts quality measurements, the Markov inequality

Table 5.2. Critical numbers $n^*_{q,M}(\alpha_q, \beta_q), n^*_{q,C}(\alpha_q, \beta_q)$, $n^*_{q,S}(\alpha_q, \beta_q)$, $n^*_q(\alpha_q, \beta_q)$ and the corresponding approximated values as functions of (α_q, β_q).

(α_q, β_q) / Critical Number	$(0.1, 0.9)$		$(0.1, 0.95)$		$(0.05, 0.9)$		$(0.05, 0.95)$	
	n^*_q	$\breve{n}^*_q$	n^*_q	$\breve{n}^*_q$	n^*_q	$\breve{n}^*_q$	n^*_q	$\breve{n}^*_q$
$n^*_{q,M}(\alpha_q, \beta_q)$	708	1519	2830	6103	2830	6103	11318	24364
$n^*_{q,C}(\alpha_q, \beta_q)$	112	239	223	480	445	959	889	1914
$n^*_{q,S}(\alpha_q, \beta_q)$	31	65	43	91	117	254	173	362
$n^*_q(\alpha_q, \beta_q)$	31	65	43	92	121	258	171	367

Table 5.3. Computation time for evaluating $n^*_{T,S}$ and $n^*_{q,S}$ of Tables 5.1-5.2.

Computation Time (sec)	$(0.1, 0.9)$	$(0.1, 0.95)$	$(0.05, 0.9)$	$(0.05, 0.95)$
$n^*_{T,S}(\alpha, \beta)$	1.2020	1.6247	5.0171	7.4426
$n^*_T(\alpha, \beta)$	0.0002	0.0002	0.0002	0.0002
$n^*_{q,S}(\alpha_q, \beta_q)$	0.1131	0.1704	0.2548	0.3593
$n^*_q(\alpha_q, \beta_q)$	0.0002	0.0002	0.0002	0.0002

and Chebyshev inequality give much larger critical numbers compared to the ones obtained by the (α, β)-Precision Theory. Specially, the above numerical illustration shows that the critical number evaluated using Markov inequality can be 20 to 70 times larger than the ones evaluated using our theory, while the ones obtained using Chebyshev inequality can be 3 to 5 times larger. As one can see, our theory gives the smallest critical numbers which guarantees the estimates precision, and can be even orders of magnitude smaller compared to numbers that classic probability inequalities can give.

As for the evaluation of critical number using simulation, the method has three disadvantages:

- no analytical property of $n^*_{T,S}$ can be investigated as a function of α and β;
- since no analytical expressions for $n^*_{T,S}$ or $n^*_{q,S}$ are available, simulations have to be repeated anew for every value of T_{up}, T_{down}, and q;
- its computation time complexity is $\mathcal{O}(Nn\log(n))$, while that of (α, β)-Precision Theory is $\mathcal{O}(1)$;
- the results are subject to randomness in the up- and downtime realizations and parts quality realizations.

CHAPTER 6

Production Systems with Cycle Overrun: Modeling, Analysis, Improvability, Bottlenecks, and Parameter Estimation

6.1 Problem Motivation

Throughput losses in production systems are usually attributed to two reasons: unreliable equipment and random part processing time (also referred to as machine cycle time). The former is typically considered in system-theoretic literature devoted to production systems, where the machine up- and downtime are assumed to be random variables (often, exponentially distributed), while the cycle time is viewed as a constant (see for instance, *Viswanadham and Narahari* (1992), *Gershwin* (1994), and *Li and Meerkov* (2009)). The latter is typically used in queuing-theoretic literature, where the processing time is assumed to be a random variable (often, also exponential), while up- and downtimes are not explicitly considered and may be viewed as "embedded" in the random processing time (see, for instance, *Askin and*

Standridge (1993), *Buzacott and Shanthikumar* (1993), *Papadopolous et al.* (1993), *Altiok* (1997), *Papadopoulos et al.* (2009), and *Curry and Feldman* (2010)).

In practice, however, one more reason for throughput losses is observed: cycle overrun. This term implies that the part processing time, τ, which is supposed to be constant, may, in fact, require additional time (i.e., overrun), τ_{OR}, leading to the total machine processing time given by

$$\tau_{total} = \tau + \tau_{OR}. \tag{6.1}$$

Such situations occur, for instance, in automated operations with a constant cycle time, τ, and manual loading/unloading operations, which may have a random component in their duration. This scenario takes place in numerous machining and welding operations. Another scenario is typical in assembly operations, where a fixed cycle time, τ, is imposed by operational conveyors, and the overrun, τ_{OR}, is enabled by push-buttons, offering the operator a possibility to occasionally stop the conveyor in order to complete the job with the desired quality. This scenario takes place, for example, in automotive paint shops and general assembly.

Two main features characterize the cycle overrun. The first one is that it may or may not take place at every cycle time; this implies that overruns occur with a certain probability. The second feature is that, given that the overrun occurs, the conditional pdf of its duration is related to the part processing time, τ. Indeed, in most cases the overrun duration is either a fraction or a small multiple of τ. These features, exacerbated by the fact that the machines with cycle overrun may have equipment-dependent up- and downtimes, make the queuing-theoretic approach inapplicable to systems with cycle overrun. The system-theoretic approach is not applicable as well: while it does consider machine reliability models in terms of up- and downtime, the cycle time is assumed to be constant.

Given that the current literature offers no analytical methods for analysis and improvement of production systems with unreliable machines and cycle overrun and taking into account that these systems are often encountered in practice, developing such methods is of importance. This is carried out in the current paper.

Specific novel results reported here are:

- A mathematical model of an unreliable machine with cycle overrun is introduced.

- A simplified version of this model is proposed, which enables analytical performance investigation.

- The relative effectiveness of a stand-alone machine throughput improvement by reducing its average downtime vs. its average overrun is investigated.

- The effect of cycle overrun on the performance of serial production line is analyzed.

- The bottleneck identification and throughput improvability in production systems with cycle overrun is investigated, and effectiveness of throughput improvement by downtime reduction vs. cycle overrun reduction is analyzed.

- The above results, obtained under the exponential assumption, are extended to non-exponential machines and non-exponential cycle overrun.

- The theoretical results derived in this paper are illustrated by a case study based on an automotive transmission case machining line.

The outline of this paper is as follows: Section 6.2 presents a mathematical model of an unreliable machine with cycle overrun. Section 6.3 develops a method for reducing this model to the one considered in the system-theoretic literature on production systems. In Section 6.4, efficacy of stand-alone machine throughput improvement by decreasing either downtime or overrun is analyzed. In Section 6.5, a method for performance analysis of serial lines with unreliable machines and cycle overrun is investigated. In Section 6.6, a bottleneck identification technique for serial lines with unreliable machines and overruns is examined, and the issues of system improvability are discussed. Section 7 extends the results obtained in Sections 6.2-6.6 under exponential assumption to production lines with non-exponential machines and non-exponential overruns. A case study is described in Section 6.8. Finally, the summary of this chapter is given in Section 6.10. The proofs are included in the Appendix E.

6.2 Mathematical Model of Unreliable Machines with Cycle Overrun

This model is defined by the following three groups of parameters/assumptions:

(a) Nominal parameters:

- *Machine cycle time* (τ) – the nominal time necessary to process a part by a machine. The term "nominal" is used to imply that the machine operates in the ideal regime, e.g., with no overruns. In large volume manufacturing, τ is practically always constant. If manual loading and unloading operations are involved, their nominal durations are included in τ.

- *Machine capacity* (c) – the nominal number of parts a stand-alone machine produces per unit of time in the ideal regime, e.g., without breakdowns and cycle overruns. If the unit of time is an hour and the cycle time is in seconds, the machine capacity is

$$c = \frac{3600}{\tau} \text{ parts/hour.} \tag{6.2}$$

(b) Reliability assumption:

- *Exponential reliability model* – machine breakdown and repair rates, λ and μ, are constant, implying that up- and downtime of the machine are distributed exponentially with parameters λ and μ, respectively. The inverses of λ and μ are average up- and downtime, T_{up} and T_{down}.

While the above machine characteristics are widely used in system-theoretic literature, the ones below are novel.

(c) Cycle overrun parameters/assumptions:

- *Overrun probability* (p_{OR}) – the probability that a cycle has an overrun. The complementary probability, $1 - p_{OR}$, is the probability that this cycle does not have an overrun.

- *Overrun distribution* ($f_{OR}^c(t)$) – the conditional pdf of the overrun duration, given that the cycle has an overrun. This distribution is assumed to be exponential with the expected value denoted as T_{OR}^c. To reflect the practical meaning of the overrun, it is assumed that T^c_{OR} is either a fraction or a small multiple of τ. Specifically, we assume that $T^c_{OR} = k_{OR}\tau$, where $k_{OR} \in (0, 2]$.

□

Based on the above, the conditional and unconditional pdf's of the overrun as well as its unconditional expected value (T_{OR}) are:

$$f^c_{OR}(t) = \frac{1}{k_{OR}\tau}\exp(-\frac{t}{k_{OR}\tau}), \quad t \geqslant 0, \tag{6.3}$$

$$f_{OR}(t) = p_{OR}\left(\frac{1}{k_{OR}\tau}\exp(-\frac{t}{k_{OR}\tau})\right) + (1 - p_{OR})\delta(t), \quad t \geqslant 0, \tag{6.4}$$

$$T_{OR} = p_{OR}k_{OR}\tau = p_{OR}T^c_{OR}, \tag{6.5}$$

where $\delta(t)$ is the Dirac delta function.

Thus, according to the above model, an exponential unreliable machine with cycle overrun is defined by five independent parameters $\{\tau, T_{up}, T_{down}, p_{OR}, k_{OR}\}$.

Note that model (a)-(c) can be extended to non-exponential distributions of up-time, downtime, and cycle overrun; some of the results in this direction are described in Section 7. Note also that, according to the above formulation, τ_{OR} in expression (6.1) is exactly T_{OR}.

6.3 Simplified Mathematical Model of Unreliable Machines with Cycle Overrun

In the system-theoretic approach, a machine is characterized by three independent parameters $\{\tau, T_{up}, T_{down}\}$. They are used to evaluate machine's efficiency, e, and stand-alone throughput, SAT, according to

$$\begin{aligned} e &= \frac{T_{up}}{T_{up} + T_{down}}, \\ SAT &= \frac{3600}{\tau} \frac{T_{up}}{T_{up} + T_{down}} \text{ parts/hour}, \end{aligned} \tag{6.6}$$

where, as in (6.2), the coefficient 3600 is used to account for τ being in seconds and SAT in parts/hour.

To simplify the description and performance analysis of a machine with cycle overrun, it is desirable to reduce the five-parameter description $\{\tau, T_{up}, T_{down}, p_{OR}, k_{OR}\}$ of a machine with cycle overrun, to a three-parameter case similar to $\{\tau, T_{up}, T_{down}\}$. This can be accomplished by embedding the unconditional duration of the overrun into T_{up} or T_{down} or τ. In this work, we use the latter due to the following fact:

Theorem 6.1. *The stand-alone throughput of an unreliable machine with cycle overrun defined by assumptions (a)-(c) is given by*

$$SAT = \frac{3600}{\tau + T_{OR}} \frac{T_{up}}{T_{up} + T_{down}} \text{ parts/hour.} \tag{6.7}$$

Proof. See Appendix E. □

Comparing (6.6) and (6.7), one can see that (6.7) corresponds to an unreliable machine with no cycle overrun, but with the cycle time $\tau + T_{OR}$ (instead of τ). In other word, from the point of stand-alone throughput, an unreliable machine with cycle overrun defined by $\{\tau, T_{up}, T_{down}, p_{OR}, k_{OR}\}$ is equivalent to the unreliable machine with the overrun duration embedded into τ and, thus, described by three parameters $\{\tau + T_{OR}, T_{up}, T_{down}\}$.

The triple $\{\tau + T_{OR}, T_{up}, T_{down}\}$ is referred to as the *simplified parametric model of unreliable machine with cycle overrun*. Sections 6.5 and 6.6 below show that this model is sufficiently precise to be used for performance analysis and improvement of production systems with machines having cycle overrun.

6.4 Improvability of Stand-alone Unreliable Machine with Cycle Overrun

Production losses of an unreliable machine with cycle overrun can be decreased by either decreasing its downtime or cycle overrun. Which one of these options is preferable?

To formalize this question, consider a simplified model of unreliable machine with cycle overrun defined by $\{\tau + T_{OR}, T_{up}, T_{down}\}$. Assume that its downtime is reduced to become rT_{down}, where the $r \in (0, 1)$ is the *reduction coefficient*. In this case, the machine is characterized by $\{\tau + T_{OR}, T_{up}, rT_{down}\}$ and referred to as *downtime-reduced machine*. Alternatively, assume that the unconditional mean of the overrun is reduced by the same fraction. This results in a machine defined by $\{\tau + rT_{OR}, T_{up}, T_{down}\}$ and is referred to as *overrun-reduced machine*. The statement below specifies which one of these machines has a larger SAT.

Theorem 6.2. *Consider a machine defined by assumptions (a)-(c). Then, for any value of* $r \in (0,1)$, *if*

$$\frac{T_{down}}{T_{up}} < p_{OR}k_{OR}, \tag{6.8}$$

the overrun-reduced machine has a larger SAT *than the downtime-reduced machine. If this inequality is reversed, the downtime-reduced machine is more productive than the overrun-reduced one.*

Proof. See Appendix E. □

Theorem 6.2 leads to the following

SAT **Improvability Indicator:**

(a) If the machine parameters are such that (6.8) holds, to improve its SAT decrease the unconditional expected values of the overrun, T_{OR}.

(b) If the machine parameters are such that (6.8) does not hold, to improve its SAT increase the machine efficiency, e.

6.5 Performance Analysis of Serial Lines with Unreliable Machines and Cycle Overrun Based on Simplified Parametric Model

While the simplified parametric model of an unreliable machine with cycle overrun precisely predicts its performance in terms of the stand-alone throughput, in a multi-machine production system with finite buffers this may not be the case. Therefore, in this section we investigate the accuracy of serial lines performance evaluation using the simplified parametric model of Section 6.3 viz-a-viz the exact stochastic model defined by assumptions (a)-(c) in Section 6.2. In addition to the throughput (TP), we investigate the accuracy of work-in-process (WIP) and probabilities of blockages (BL) and starvations (ST). First we consider the case of two-machine lines (where closed formulas for all performance metrics are available in *Li and Meerkov* (2009)) and then address the general case of ($M > 2$)-machines (using the aggregation procedure of *Bai et al.* (2021)). In both cases, the accuracy of using the simplified model is evaluated in terms of the errors defined by:

$$
\begin{aligned}
\epsilon_{TP} &= \frac{|TP_{sim} - TP|}{TP_{sim}} \cdot 100\%, \\
\epsilon_{WIP} &= \frac{1}{M-1} \sum_{i=1}^{M-1} \frac{|WIP_{i,sim} - WIP_i|}{N_i} \cdot 100\%, \\
\epsilon_{BL} &= \frac{1}{M-1} \sum_{i=1}^{M-1} |BL_{i,sim} - BL_i|, \\
\epsilon_{ST} &= \frac{1}{M-1} \sum_{i=2}^{M} |ST_{i,sim} - ST_i|,
\end{aligned}
\tag{6.9}
$$

where the symbols with subscript 'sim' refer to the performance metrics evaluated by simulations, and the symbols without the subscript refer to the same performance metrics calculated analytically (either by closed formulas or by aggregation). Note that in the case of two-machine lines the summation signs in (6.9) are omitted.

6.5.1 Two-machine Lines

In the two-machine case, we generate 100 lines, with machine and buffer parameters selected randomly and equiprobably from the following sets:

$$T_{down,i} \in [3, 10], e_i \in [0.6, 0.95], c_i \in [1, 2], p_{OR,i} \in [0, 1], k_{OR,i} \in [0.2, 2], i = 1, 2;$$
$$N = \lceil h \max(c_1 T_{down,1}, c_2 T_{down,2}) \rceil, h \in [2, 4], \tag{6.10}$$

where h denotes the level of buffering, which protects a machine against job losses during the adjacent machine's downtime.

The analytical calculations for each of these lines have been carried out using expressions (11.13)-(11.17) of *Li and Meerkov* (2009). The results obtained are summarized in Tables 6.1. Based on these results, we conclude that the simplified parametric machine model is acceptable for two-machine systems evaluation.

Table 6.1. Accuracy of performance metrics evaluation using the simplified parametric model in two-machine lines.

(a) Accuracy of TP evaluation

	Mean value
TP_{sim}	40.7941
TP	40.8795
ϵ_{TP}	0.29%

(b) Accuracy of WIP evaluation

	Mean value
WIP_{sim}	17.8947
WIP	17.8821
ϵ_{WIP}	1.72%

(c) Accuracy of BL evaluation

	Mean value
$BL_{1,sim}$	0.1140
BL_1	0.1120
ϵ_{BL}	0.0021

(d) Accuracy of ST evaluation

	Mean value
$ST_{2,sim}$	0.1140
ST_2	0.1122
ϵ_{ST}	0.0021

6.5.2 ($M > 2$)-machine Lines

In the ($M > 2$) case, we consider $M \in \{3, 5, 10, 15, 20\}$, and for each M generate 100 lines, with machine and buffer parameters selected randomly and equiprobably from the following sets:

$$T_{down,i} \in [3, 10], e_i \in [0.6, 0.95], c_i \in [1, 2], p_{OR,i} \in [0, 1], k_{OR,i} \in [0.2, 2], i = 1, \ldots, M;$$
$$N_j = \lceil h_j \max(c_j T_{down,j}, c_{j+1} T_{down,j+1}) \rceil, h_j \in [2, 4], j = 1, \ldots, M - 1. \tag{6.11}$$

The results obtained are presented in Table 6.2. From these results, we observe that the accuracy of performance metrics evaluation is decreasing as a function of M, and for large M the errors become relatively large. This is because in the former case the errors are not only due to the reduction of the exact model to a simplified one, but also due to the errors inherent in the aggregation procedure of *Bai et al.* (2021). Nevertheless, since in most practical cases the data of machine parameters is rarely available with high precision, we conclude that the simplified machine reliability model is still acceptable in most practical systems evaluation with $M \leqslant 20$.

Table 6.2. Accuracy of performance metrics evaluation using the simplified parametric model in $(M > 2)$-machine lines.

(a) Accuracy of TP evaluation

Mean value	$M = 3$	$M = 5$	$M = 10$	$M = 15$	$M = 20$
TP_{sim}	34.374	30.426	26.514	23.906	23.262
TP	34.414	30.484	26.797	24.588	24.119
ϵ_{TP}	0.39%	0.53%	1.44%	3.32%	4.26%

(b) Accuracy of WIP evaluation

Mean value	$M = 3$	$M = 5$	$M = 10$	$M = 15$	$M = 20$
$\frac{1}{M-1}\sum_{i=1}^{M-1} WIP_{i,sim}$	19.209	18.609	16.856	17.315	17.970
$\frac{1}{M-1}\sum_{i=1}^{M-1} WIP_{i}$	19.164	18.904	18.026	21.433	23.091
ϵ_{WIP}	2.08%	2.13%	4.80%	11.68%	14.37%

(c) Accuracy of BL evaluation

Mean value	$M = 3$	$M = 5$	$M = 10$	$M = 15$	$M = 20$
$\frac{1}{M-1}\sum_{i=1}^{M-1} BL_{i,sim}$	0.1620	0.1661	0.1630	0.1811	0.1869
$\frac{1}{M-1}\sum_{i=1}^{M-1} BL_{i}$	0.1627	0.1681	0.1765	0.2254	0.2436
ϵ_{BL}	0.0041	0.0058	0.0163	0.0452	0.0564

(d) Accuracy of ST evaluation

Mean value	$M = 3$	$M = 5$	$M = 10$	$M = 15$	$M = 20$
$\frac{1}{M-1}\sum_{i=2}^{M} ST_{i,sim}$	0.1234	0.1509	0.1841	0.1973	0.1997
$\frac{1}{M-1}\sum_{i=2}^{M} ST_{i}$	0.1231	0.1501	0.1697	0.1447	0.1319
ϵ_{ST}	0.0036	0.0050	0.0185	0.0546	0.0686

6.6 Bottleneck Identification and Improvability of Serial Lines with Unreliable Machines and Cycle Overrun

6.6.1 Bottleneck Identification

The notion of bottlenecks in serial lines is formulated in *Li and Meerkov* (2009) as follows:

Definition 6.1. Machine m_i is the bottleneck (BN) of a serial line with M unreliable machines if

$$\frac{\partial TP}{\partial c_i} > \frac{\partial TP}{\partial c_j}, \ \forall j \neq i, \tag{6.12}$$

where c_k is the capacity of the k-th machine, $k = 1, \ldots, M$.

Since evaluating analytically the partial derivatives involved in (6.12) is all but impossible, *Li and Meerkov* (2009) provide the following

BN Identification Procedure:

- Evaluate BL and ST of all machines in the system (either by calculation or by measurements on the factory floor.)
- Assign arrows in each pair of consecutive machines according to the rule:
  - if $BL_i > ST_{i+1}$, assign the arrow pointing from m_i to m_{i+1};
  - if $BL_i < ST_{i+1}$, assign the arrow pointing from m_{i+1} to m_i;
- Then,

- if there is a single machine with no emanating arrows, it is the BN in the sense of (6.12);
- if there are multiple machines with no emanating arrows, each of them is referred to as a Local BN (L-BN), with its severity defined as

$$S_1 = |ST_2 - BL_1|, \text{ if L-BN} = m_1,$$

$$S_M = |ST_M - BL_{M-1}|, \text{ if L-BN} = m_M,$$

$$S_i = |ST_{i+1} - BL_i| + |ST_i - BT_{i-1}|, \text{ if L-BN} = m_i, i = 2, \ldots, M-1;$$

- L-BN with the largest severity is the Primary BN (P-BN).

It is shown in *Li and Meerkov* (2009) that this identification procedure determines the BN in systems with machines having no overruns with high accuracy. In this subsection, we verify by simulations whether this approach works for serial lines with machines having overruns, modeled by the simplified parametric model. This is carried out as follows:

We consider $(M > 2)$-machine serial lines with cycle overrun for $M \in \{3, 5, 10, 15, 20\}$. For each M, we generate 100 lines, with parameters selected randomly and equiprobably from the sets defined in (6.11). For each of the 500 lines, we identify BN(s) using BN Identification Procedure and assess its accuracy by comparing the results with those identified using numerical evaluation of the derivatives involved in (6.12). As it turns out, in all 500 lines considered, the BN identified numerically is one of the L-BNs identified by BN Identification Procedure. Table 6.3 shows the number of lines, where P-BN identified by this procedure is the same as the one identified numerically.

Based on the above, we conclude that the BN Identification Procedure has an acceptable accuracy in serial lines with cycle overrun modeled by the simplified parametric model.

Table 6.3. The number of lines where P-BN identified by BN Identification Procedure is the same as the one identified numerically.

$M = 3$	$M = 5$	$M = 10$	$M = 15$	$M = 20$
100	97	91	90	89

6.6.2 Throughput Improvability

Section 6.4 provides the SAT Improvability Indicator for a single machine with overruns. Can this indicator be used for improving serial lines with unreliable machines and cycle overruns? We answer this question by simulations.

In the above-mentioned 500 lines, we identify BN or P-BN, using the BN Identification Procedure, and reduce its T_{down} or T_{OR} by 10%, using SAT Improvability Indicator. Table 6.4 shows the number of lines for each M, where SAT Improvability Indicator turns out to be optimal in the sense that improving the BN or P-BN in accordance with SAT Improvability Indicator leads to the largest system throughput improvement, as compared with all other ways of BN or P-BN modification. Based on this fact, we conclude that SAT Improvability Indicator can be used for improvement of serial lines with overruns described by the simplified parametric model of the machines.

Using the results of Sections 6.4 and 6.5, as well as Subsection 6.6.1, we formulate the following

Continuous Improvement Design Procedure:

Table 6.4. The number of lines, where SAT Improvability Indicator leads to the largest serial line improvement.

$M = 3$	$M = 5$	$M = 10$	$M = 15$	$M = 20$
98	97	95	93	94

(a) Using the BN Identification Procedure, determine the BN or P-BN.

(b) Using the SAT Improvability Indicator, decrease either T_{down} or T_{OR} of this machine.

(c) Calculate the throughput of the improved system, TP_{imp}.

(d) If $TP_{imp} > TP_{des}$, where TP_{des} is the desired throughput, stop; else go to (a).

6.7 Improvability of Serial Lines with Non-exponential Machines and Non-exponential Cycle Overrun

6.7.1 Preliminaries

The results reported in Sections 6.3-6.6 are obtained under the exponential assumption on machines' reliability and overruns. In practice, however, this assumption may not hold. So, can the obtained results be applied for designing continuous improvement projects in practice? This section provides a positive answer to this question.

As one can see, the proofs of Theorems 6.1 and 6.2 do not rely on the exponential assumption and, thus, hold for non-exponential machines and overruns as well.

Unfortunately, this is not the case for the results of Sections 6.5 and 6.6, where TP is evaluated assuming that the machines and overruns are exponential. Thus, an additional investigation is necessary.

The investigation described below is conceptually related to several facts concerning serial lines with non-exponential machines, but with no overruns. These facts take place assuming that the breakdown and repair rates of machines, $\lambda(t)$ and $\mu(t)$, are increasing functions of time. This implies that the probability of machine breakdown or repair in an infinitesimal time interval $(t, t + \delta t)$ is increasing in t (which is a practically plausible case). Under this assumption, the following facts take place:

- The distributions of up- and downtime, induced by these $\lambda(t)$ and $\mu(t)$, have coefficients of variation, CV, less than 1. This is proved analytically in *Barlow and Proschan* (1996) and *Li and Meerkov* (2009). An empirical study (*Inman* (1999)) shows that in most cases manufacturing equipment on the factory floor indeed has $CV < 1$.

- The throughput of serial lines with machines having $CV_{up} < 1$ and $CV_{down} < 1$ is a monotonically decreasing function of these CV's on the interval $[0, 1]$. This is shown by simulations in *Li and Meerkov* (2009) for Weibull, gamma, and log-normal distributions and hypothesized that it holds for any distribution of up- and downtime with $CV < 1$.

- Thus, the throughput of a serial line with exponential machines, (i.e., the machines having $CV_{up} = 1$ and $CV_{down} = 1$, is the lower bound of the throughput of lines with non-exponential machines having $CV_{up} < 1$ and $CV_{down} < 1$.

- Hence, if a continuous improvement project is designed to achieve the desired throughput under the exponential assumption, the throughput of the real system with non-exponential machines, will be at least as high as that of the exponential one. Therefore, the goal of the improvement project will be achieved (or, perhaps, "over-achieved").

It is shown below by simulations that similar facts take place for serial lines with machines having cycle overrun as well. First, we investigate the case of exponential machines and non-exponential overruns and then the general case, where both machines and overruns are non-exponential.

6.7.2 Serial Lines with Exponential Machines and Non-exponential Overruns

We carry out the simulations for $M \in \{5, 10\}$ and generate 50 lines with exponential machines having non-exponential overruns, where machine and buffer parameters are selected randomly and equiprobably from the sets defined in (6.11). The conditional distribution of each machine's overrun is selected randomly and equiprobably from the set {Weibull, gamma, log-normal}, with pdf's given by:

$$\begin{aligned}
&\text{Weibull:} && f^c_{OR}(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} e^{-(\frac{t}{\alpha})^{\beta}}, && t \geqslant 0,\\
&\text{gamma:} && f^c_{OR}(t) = \frac{1}{\Gamma(\beta)\alpha^{\beta}} t^{\beta-1} e^{-\frac{t}{\alpha}}, && t \geqslant 0,\\
&\text{log-normal:} && f^c_{OR}(t) = \frac{1}{\sqrt{2\pi}\beta t} e^{-\frac{(\ln t - \frac{1}{\alpha})^2}{2\beta^2}}, && t \geqslant 0,
\end{aligned} \tag{6.13}$$

where

$$\Gamma(x) = \int_0^{\infty} s^{x-1} e^{-s} ds \tag{6.14}$$

and α and β are the parameters to be selected so that their expected values are $k_{OR}\tau$ (similar to (6.3) for the exponential distribution).

This can be accomplished using two equations – one based on the expected value and the second on the coefficient of variation of distributions (6.13). For instance, in the case of Weibull distribution these equations are:

$$T_{OR}^c = \alpha\Gamma(1 + \frac{1}{\beta}), \tag{6.15}$$

$$CV = \sqrt{\frac{\Gamma(1 + \frac{2}{\beta})}{\Gamma^2(1 + \frac{1}{\beta})} - 1}\,, \tag{6.16}$$

where $T_{OR}^c = k_{OR}\tau$ and $CV < 1$. If these equations have a unique solution (α, β), the required conditional distribution of overruns is fully defined.

Theorem 6.3. *For any $T_{OR}^c > 0$ and $CV < 1$, equations* (6.15) *and* (6.16) *have a unique solution.*

Proof. See Appendix E. □

Thus, for a given $CV < 1$, a unique β can be found from (6.16) and then for a given T_{OR}^c a unique α from (6.15). For example, if T_{OR}^c = 30sec and $CV = 0.75$, then $\alpha = 32.7053$ and $\beta = 1.3476$. Parameters α and β for gamma and log-normal distributions can be calculated in a similar manner (using the formulas for their mean and coefficient of variation in the right-hand sides of (6.15) and (6.16)).

For each of the conditional distributions, thus obtained, and each CV in the set $\{0.1, 0.25, 0.5, 0.75, 1\}$, we calculate β using (6.16) and, for selected T_{OR}^c, calculate α using (6.15). With these parameters, we simulate the 50 lines mentioned above.

The obtained results are compared with throughput of the lines having exponential machines and exponential overruns with the same uptime, downtime, and unconditional mean of cycle overrun. The average value of the former over 50 lines is denoted as TP_{NEOR} and the latter as TP_E. The difference between TP_{NEOR} and TP_E is evaluated by

$$\zeta = \frac{TP_{NEOR} - TP_E}{TP_E} 100\%. \tag{6.17}$$

The behavior of ζ as a function of CV for $M \in \{5, 10\}$ is shown in Table 6.5. As one can see, the difference is practically negligible. Thus, the effect of non-exponential overruns in the case of Weibull, gamma, and log-normal distributions is insignificant. We hypothesize that the same effect takes place for any overrun distribution with $CV < 1$.

Table 6.5. The behavior of ζ as a function of CV for $M \in \{5, 10\}$.

CV	$M = 5$	$M = 10$
0.1	0.4266%	0.4366%
0.25	0.4151%	0.4284%
0.5	0.3682%	0.3877%
0.75	0.1427%	0.1791%
1	0.0063%	0.0054%

6.7.3 Serial Lines with Non-exponential Machines and Non-exponential Overruns

In this subsection, we consider the lines discussed in Subsection 6.7.2, but with non-exponential reliability models selected randomly and equiprobably from the set {Weibull, gamma, log-normal}. For $CV \in \{0.1, 0.25, 0.5, 0.75, 1\}$, we evaluate the lines throughput using simulations. The average value of the throughput over the 50 lines is denoted as TP_{NE}. Figure 6.1 shows the behavior of TP_{NE} as a function of $CV < 1$ for $M \in \{5, 10\}$. As one can see, this behavior is similar to that of non-exponential machines having no overruns: it is a monotonically decreasing function with the minimum at $CV = 1$.

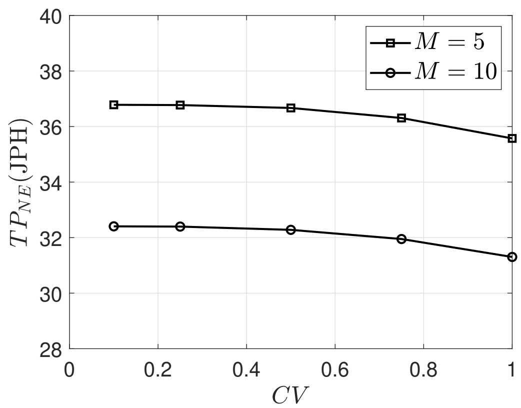


Figure 6.1. The behavior of TP_{NE} as a function of CV for $M \in \{5, 10\}$.

Thus, TP of serial lines with exponential machines and exponential overruns is the lower bound of serial lines with non-exponential machines and non-exponential overruns for the case of Weibull, gamma and log-normal distributions with $CV < 1$. Hence, the former can be used to design continuous improvement projects for the latter: the improved system would exhibit the performance at least as good as that predicted under the exponential assumption. We hypothesize that this conclusion takes place for any machine reliability model and cycle overrun if $CV < 1$.

The case study described in the next section is based on this hypothesis.

6.8 Case Study

6.8.1 Preliminaries

In this case study, we consider a production system motivated by an automotive transmission case machining line. For confidentiality reasons, all machine and buffer parameters have been modified, preserving, however, qualitative features of the system performance. Due to these modifications, the system throughput, which has been measured on the factory floor, could not be used for validating the mathematical model of this system. Therefore, we have created a computer simulation model of the system at hand, populated it by the modified data, and used it as the "real system" for model "validation".

6.8.2 Mathematical Modeling

The system at hand has been modeled as a serial line consisting of 12 operations separated by finite buffers (see Figure 6.2, where the numbers in the rectangles represent the modified buffers capacity). Based on the conclusion of the previous section, we assume that the machines and cycle overruns, if any, are exponential.

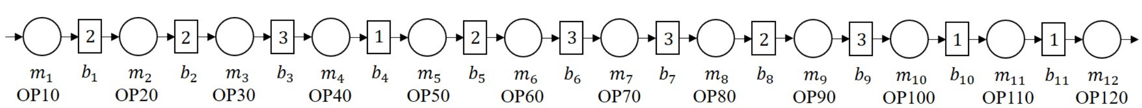


Figure 6.2. Structural model of the modified production line.

The parametric model of this system has been constructed using eight weeks of modified data. For each week, the data provide machines' τ, T_{up}, T_{down}, p_{OR}, and

k_{OR}. Based on these data, e and T_{OR} are calculated and included in Table 6.6.

Using these data and the aggregation procedure of *Bai et al.* (2021), we calculated the system throughput for all eight weeks (denoted as TP) and compared it with that evaluated by simulations using the digital twin populated by the respective week data (denoted as TP_{sim}). The results, shown in Table 6.7, indicate that the mathematical model is "validated".

Table 6.6. Weekly raw data.

(a) Week 1.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	8.3406	3.024	0.7339	0.3563	0.6215	26.5723
OP20	119	24.9552	4.5164	0.8468	0	0	0
OP30	120	49.9567	4.1475	0.9233	0.3592	0.1651	7.1172
OP40	120	19.9509	3.3174	0.8574	0.1644	0.0994	1.9621
OP50	106	50.0429	3.6922	0.9313	0	0	0
OP60	120	16.6674	3.5635	0.8239	0.3318	0.2778	11.0585
OP70	120	49.9549	3.4369	0.9356	0.2762	0.3389	11.2307
OP80	120	8.3601	2.57	0.7649	0.2529	0.0108	0.3268
OP90	120	24.9889	2.826	0.8984	0	0	0
OP100	113	14.286	3.0102	0.826	0	0	0
OP110	120	19.9701	2.4306	0.8915	0.3136	0.1366	5.1407
OP120	105	50.0198	4.1669	0.9231	0	0	0

(b) Week 2.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	16.683	3.5742	0.8236	0.6169	0.6012	44.5056
OP20	119	33.3561	4.3471	0.8847	0	0	0
OP30	120	16.6831	3.5722	0.8236	0.3214	0.1227	4.7307
OP40	120	16.6979	3.3335	0.8336	0.1818	0.4583	9.9947
OP50	106	49.9756	2.3848	0.9545	0	0	0
OP60	120	33.3354	3.4229	0.9069	0.3356	0.4993	20.1068
OP70	120	100.0144	3.7587	0.9638	0.2784	0.37	12.3582
OP80	120	20.0006	2.1139	0.9044	0.2557	0.0135	0.4154
OP90	120	25.0173	2.8394	0.8981	0	0	0
OP100	113	25.0066	3.5459	0.8758	0	0	0
OP110	120	25.013	2.1959	0.9193	0.2976	0.4032	14.3994
OP120	105	50.0034	4.1134	0.924	0	0	0

(c) Week 3.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	8.3052	3.109	0.7276	0.417	0.5691	28.4766
OP20	119	16.6606	3.6998	0.8183	0	0	0
OP30	120	19.995	3.3221	0.8575	0.3581	0.141	6.0611
OP40	120	20.0183	3.1142	0.8654	0.152	0.1147	2.0931
OP50	106	19.9967	4.3187	0.8224	0	0	0
OP60	120	25.0236	3.8374	0.867	0.4213	0.38	19.2123
OP70	120	20.0048	3.4488	0.853	0.266	0.4015	12.8192
OP80	120	20.0148	3.6823	0.8446	0.2989	0.1561	5.5977
OP90	120	33.3934	3.1262	0.9144	0	0	0
OP100	113	20.0199	3.3236	0.8576	0	0	0
OP110	120	24.9656	2.6729	0.9033	0.3389	0.1745	7.0953
OP120	105	33.397	6.6547	0.8338	0	0	0

(d) Week 4.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	7.6206	3.1877	0.7051	0.3697	0.6626	29.3931
OP20	119	6.5843	2.8958	0.6945	0	0	0
OP30	120	19.999	2.8337	0.8759	0.3317	0.1905	7.5803
OP40	120	20.0329	2.6112	0.8847	0.1607	0.1539	2.9673
OP50	106	11.1021	2.9251	0.7915	0	0	0
OP60	120	19.9859	2.8353	0.8758	0.4195	0.2644	13.3123
OP70	120	50.0824	3.2984	0.9382	0.2298	0.5568	15.3532
OP80	120	20.0935	3.7012	0.8445	0.3224	0.4793	18.5469
OP90	120	33.3307	3.1948	0.9125	0	0	0
OP100	113	20.0855	3.321	0.8581	0	0	0
OP110	120	33.3265	2.6876	0.9254	0.3454	0.192	7.9598
OP120	105	99.9323	5.8691	0.9445	0	0	0

(e) Week 5.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	14.2714	4.7327	0.751	0.4465	0.5577	29.8829
OP20	119	7.6328	3.6655	0.6756	0	0	0
OP30	120	8.3027	1.7795	0.8235	0.3367	0.1548	6.2526
OP40	120	14.2686	4.7353	0.7508	0.1717	0.1309	2.6987
OP50	106	16.6899	2.7667	0.8578	0	0	0
OP60	120	24.9749	3.9629	0.8631	0.4232	0.5553	28.197
OP70	120	49.9563	5.252	0.9049	0.2733	0.4473	14.6694
OP80	120	16.7332	4.001	0.807	0.259	0.0371	1.1517
OP90	120	33.3859	4.1533	0.8894	0	0	0
OP100	113	25.0217	3.8278	0.8673	0	0	0
OP110	120	25.0765	4.1712	0.8574	0.323	0.1853	7.183
OP120	105	99.9755	7.6868	0.9286	0	0	0

(f) Week 6.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	10.0586	3.7924	0.7262	0.4133	0.4509	22.3628
OP20	119	14.3617	3.3406	0.8113	0	0	0
OP30	120	5.8879	3.6987	0.6142	0.3322	0.0682	2.7178
OP40	120	12.4703	3.3201	0.7897	0.1734	0.1725	3.5902
OP50	106	20.0145	2.8557	0.8751	0	0	0
OP60	120	25.0407	3.8424	0.867	0.3954	0.5403	25.6384
OP70	120	50.0236	3.5653	0.9335	0.2965	0.4262	15.1605
OP80	120	20.0894	3.0354	0.8687	0.3154	0.3272	12.3836
OP90	120	33.2873	4.9632	0.8702	0	0	0
OP100	113	12.4909	2.7645	0.8188	0	0	0
OP110	120	33.3266	2.9013	0.9199	0.3151	0.2117	8.0047
OP120	105	99.9662	7.6247	0.9291	0	0	0

(g) Week 7.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	6.6949	2.2091	0.7519	0.5653	0.9904	67.1901
OP20	119	12.5163	3.5632	0.7784	0	0	0
OP30	120	5.8558	2.2167	0.7254	0.3218	0.1063	4.1054
OP40	120	16.6304	3.0973	0.843	0.1695	0.1131	2.2999
OP50	106	33.2742	3.4126	0.907	0	0	0
OP60	120	24.9705	3.8251	0.8672	0.3981	0.4726	22.5791
OP70	120	99.9469	3.1556	0.9694	0.3125	0.3861	14.4802
OP80	120	14.2309	2.9193	0.8298	0.2843	0.1452	4.9556
OP90	120	24.9441	3.1127	0.8891	0	0	0
OP100	113	24.9722	3.6785	0.8716	0	0	0
OP110	120	33.3025	3.8017	0.8975	0.3	0.1983	7.1391
OP120	105	100.0185	9.001	0.9174	0	0	0

(h) Week 8.

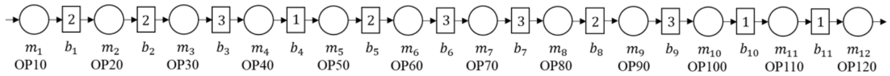
	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	7.7444	2.7936	0.7349	0.3993	0.6681	32.0122
OP20	119	19.9792	7.6559	0.723	0	0	0
OP30	120	7.136	2.4903	0.7413	0.3354	0.0917	3.6886
OP40	120	33.3248	3.8121	0.8974	0.1654	0.1311	2.6033
OP50	106	33.2792	2.8491	0.9211	0	0	0
OP60	120	14.3036	2.2152	0.8659	0.5069	0.9835	59.8269
OP70	120	33.3474	6.6485	0.8338	0.2851	0.3096	10.5952
OP80	120	20.0423	3.4521	0.8531	0.2603	0.1068	3.3357
OP90	120	25.0204	4.3444	0.8521	0	0	0
OP100	113	50.0636	3.6599	0.9319	0	0	0
OP110	120	16.6557	3.1142	0.8425	0.2919	0.1568	5.4932
OP120	105	100.0047	8.3081	0.9233	0	0	0

Table 6.7. Model validation.

Throughput (JPH)	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
TP_{sim}	16.4970	16.9290	15.7003	15.2434	14.5075	14.4710	13.8117	13.8078
TP	16.6910	17.0372	16.0725	15.3409	14.9029	14.6863	13.7300	14.1172
ϵ_{TP}	1.18%	0.64%	2.37%	0.64%	2.73%	1.49%	0.59%	2.24%

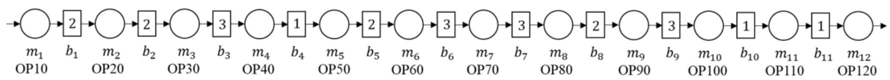
6.8.3 Weekly Performance Analysis

The performance analysis has been carried out based on the weekly data and the aggregation procedure developed in *Bai et al.* (2021). The results are shown in Figure 6.3 for Weeks 1-8.



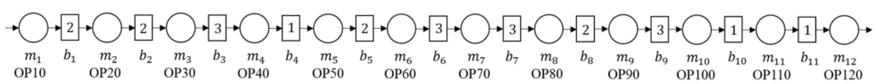
c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.73	0.85	0.92	0.86	0.93	0.82	0.94	0.76	0.90	0.83	0.89	0.92
T_{OR} (sec)	26.57	0.00	7.12	1.96	0.00	11.06	11.23	0.33	0.00	0.00	5.14	0.00
SAT (JPH)	18.03	25.62	26.15	25.31	31.63	22.63	25.67	22.88	26.95	26.31	25.65	31.65
WIP	0.43	0.65	1.08	0.28	1.05	1.12	0.48	0.22	0.46	0.42	0.18	
ST	0.00	0.27	0.27	0.20	0.33	0.16	0.19	0.16	0.31	0.28	0.31	0.44
BL	0.05	0.04	0.08	0.12	0.19	0.07	0.18	0.06	0.06	0.03	0.01	0.00
S	0.22					0.15		0.27				
TP	16.6910JPH											

(a) Week 1.



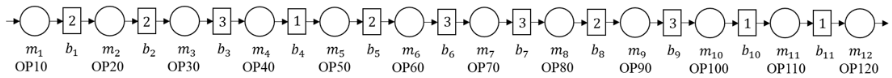
c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.82	0.88	0.82	0.83	0.95	0.91	0.96	0.90	0.90	0.88	0.92	0.92
T_{OR} (sec)	44.51	0.00	4.73	9.99	0.00	20.11	12.36	0.42	0.00	0.00	14.40	0.00
SAT (JPH)	18.02	26.76	23.77	23.08	32.42	23.30	26.21	27.04	26.94	27.90	24.62	31.68
WIP	0.35	0.82	0.85	0.09	0.55	0.41	0.22	0.19	0.34	0.53	0.17	
ST	0.00	0.29	0.20	0.20	0.43	0.23	0.29	0.29	0.30	0.32	0.28	0.43
BL	0.04	0.05	0.04	0.03	0.06	0.02	0.06	0.06	0.04	0.04	0.01	0.00
TP	17.0372JPH											

(b) Week 2.



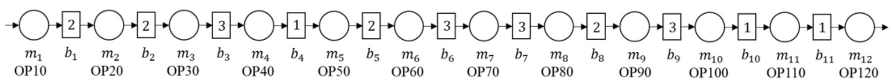
c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.73	0.82	0.86	0.87	0.82	0.87	0.85	0.84	0.91	0.86	0.90	0.83
T_{OR} (sec)	28.48	0.00	6.06	2.09	0.00	19.21	12.82	5.60	0.00	0.00	7.10	0.00
SAT (JPH)	17.64	24.75	24.49	25.52	27.93	22.42	23.12	24.21	27.43	27.32	25.59	28.59
WIP	0.53	0.95	1.26	0.40	1.03	1.14	0.31	0.18	0.39	0.47	0.44	
ST	0.00	0.24	0.21	0.17	0.24	0.19	0.17	0.25	0.35	0.33	0.31	0.37
BL	0.06	0.06	0.10	0.18	0.17	0.08	0.11	0.06	0.05	0.05	0.05	0.00
TP	16.0725JPH											

(c) Week 3.



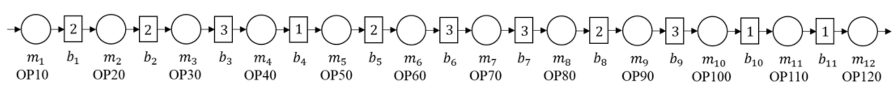
c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.71	0.69	0.88	0.88	0.79	0.88	0.94	0.84	0.91	0.86	0.93	0.94
T_{OR} (sec)	29.39	0.00	7.58	2.97	0.00	13.31	15.35	18.55	0.00	0.00	7.96	0.00
SAT (JPH)	16.99	21.01	24.72	25.90	26.88	23.65	24.95	21.94	27.38	27.34	26.03	32.38
WIP	0.69	0.47	0.72	0.28	0.59	0.73	0.39	0.14	0.29	0.24	0.13	
ST	0.00	0.18	0.31	0.28	0.31	0.28	0.27	0.22	0.38	0.37	0.37	0.50
BL	0.07	0.01	0.04	0.12	0.06	0.04	0.13	0.05	0.04	0.02	0.01	0.00
TP	15.3409JPH											

(d) Week 4.



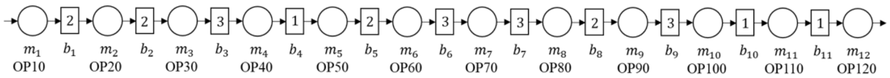
c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.75	0.68	0.82	0.75	0.86	0.86	0.90	0.81	0.89	0.87	0.86	0.93
T_{OR} (sec)	29.88	0.00	6.25	2.70	0.00	28.20	14.67	1.15	0.00	0.00	7.18	0.00
SAT (JPH)	18.04	20.44	23.48	22.03	29.13	20.97	24.19	23.98	26.68	27.63	24.27	31.84
WIP	0.89	1.00	1.30	0.32	0.96	0.97	0.34	0.20	0.40	0.48	0.17	
ST	0.00	0.14	0.20	0.15	0.33	0.19	0.23	0.25	0.36	0.37	0.32	0.50
BL	0.12	0.04	0.12	0.11	0.16	0.08	0.16	0.08	0.06	0.07	0.02	0.00
TP	14.9029JPH											

(e) Week 5.



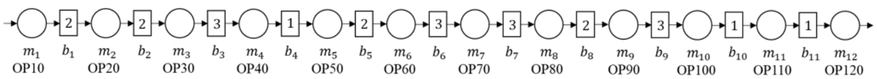
c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.73	0.81	0.61	0.79	0.88	0.87	0.93	0.87	0.87	0.82	0.92	0.93
T_{OR} (sec)	22.36	0.00	2.72	3.59	0.00	25.64	15.16	12.38	0.00	0.00	8.00	0.00
SAT (JPH)	18.36	24.54	18.02	23.00	29.72	21.43	24.86	23.62	26.11	26.09	25.87	31.86
WIP	0.96	1.87	0.93	0.21	0.73	0.63	0.33	0.19	0.39	0.26	0.16	
ST	0.00	0.16	0.07	0.23	0.39	0.25	0.30	0.27	0.36	0.35	0.39	0.50
BL	0.14	0.20	0.05	0.08	0.11	0.03	0.12	0.09	0.05	0.02	0.02	0.00
S	0.02		0.31									
TP	14.6863JPH											

(f) Week 6.



c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.75	0.78	0.73	0.84	0.91	0.87	0.97	0.83	0.89	0.87	0.90	0.92
T_{OR} (sec)	67.19	0.00	4.11	2.30	0.00	22.58	14.48	4.96	0.00	0.00	7.14	0.00
SAT (JPH)	14.46	23.55	21.04	24.81	30.80	21.90	25.95	23.91	26.67	27.77	25.41	31.46
WIP	0.38	0.65	0.47	0.13	0.57	0.40	0.26	0.15	0.28	0.32	0.19	
ST	0.00	0.31	0.24	0.34	0.47	0.31	0.39	0.32	0.41	0.42	0.40	0.52
BL	0.03	0.02	0.02	0.06	0.08	0.02	0.11	0.06	0.04	0.04	0.03	0.00
TP						13.7300JPH						

(g) Week 7.



c (JPH)	30.00	30.25	30.00	30.00	33.96	30.00	30.00	30.00	30.00	31.86	30.00	34.29
e	0.73	0.72	0.74	0.90	0.92	0.87	0.83	0.85	0.85	0.93	0.84	0.92
T_{OR} (sec)	32.01	0.00	3.69	2.60	0.00	59.83	10.60	3.34	0.00	0.00	5.49	0.00
SAT (JPH)	17.40	21.87	21.58	26.35	31.28	17.33	22.98	24.90	25.56	29.69	24.17	31.66
WIP	0.75	1.04	0.71	0.46	1.38	0.83	0.25	0.18	0.20	0.40	0.18	
ST	0.00	0.21	0.23	0.30	0.28	0.09	0.25	0.32	0.36	0.46	0.34	0.51
BL	0.14	0.06	0.04	0.17	0.32	0.08	0.10	0.08	0.03	0.05	0.02	0.00
S	0.07					0.34						
TP						14.1172JPH						

(h) Week 8.

Figure 6.3. Performance analysis based on weekly data.

6.8.4 Data for Continuous Improvement Project

The weekly data exhibit substantial variability. This is obvious from Table 6.6 and is supported by the coefficients of variation of e and T_{OR} calculated using all eight weeks data and shown in the first columns of Table 6.8. This observation makes it necessary to "pre-process" the data in order to decrease its variability and use the less variable data for the continuous improvement project design. This is accomplished by averaging machine parameters based on two or four weeks' data. The resulting coefficients of variation are shown in the second and third columns of Table 6.8.

Clearly, averaging over four weeks results in a relatively low variability of machine parameters, and this data (shown in Table 6.9) is used for continuous improvement project design. (Note that averaging over all eight weeks, i.e., over two months, might not be desirable, since machine parameters are typically non-stationary and evolve in time.)

Table 6.8. Coefficients of variation the original and averaged data.

(a) Data for e

	CV_e	$CV_{e,2}$	$CV_{e,4}$
OP10	0.0444	0.0300	0.0044
OP20	0.0900	0.0645	0.0411
OP30	0.1163	0.0905	0.0902
OP40	0.0545	0.0499	0.0238
OP50	0.0597	0.0584	0.0087
OP60	0.0243	0.0029	0.0015
OP70	0.0506	0.0230	0.0067
OP80	0.0462	0.0044	< 0.0001
OP90	0.0220	0.0186	0.0172
OP100	0.0375	0.0264	0.0104
OP110	0.0316	0.0189	0.0171
OP120	0.0347	0.0169	0.0100

(b) Data for T_{OR}

	$CV_{T_{OR}}$	$CV_{T_{OR},2}$	$CV_{T_{OR},4}$
OP10	0.3863	0.2588	0.0802
OP30	0.3068	0.2186	0.2065
OP40	0.7068	0.4087	0.2065
OP60	0.569	0.4155	0.3629
OP70	0.129	0.0924	0.0295
OP80	1.0368	0.729	0.0655
OP110	0.3431	0.1596	0.1085

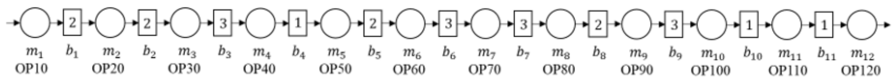
Table 6.9. Averaged data for Weeks 5-8.

	τ (sec)	T_{up} (min)	T_{down} (min)	e	p_{OR}	k_{OR}	T_{OR} (sec)
OP10	120	9.6923	3.382	0.7413	0.4561	0.6668	36.495
OP20	119	13.6225	4.5563	0.7494	0	0	0
OP30	120	6.7956	2.5463	0.7274	0.3315	0.1052	4.1861
OP40	120	19.1735	3.7412	0.8367	0.17	0.1369	2.7935
OP50	106	25.8145	2.971	0.8968	0	0	0
OP60	120	22.3224	3.4614	0.8658	0.4309	0.6379	32.9867
OP70	120	58.3186	4.6553	0.9261	0.2918	0.3923	13.7395
OP80	120	17.774	3.352	0.8413	0.2798	0.1541	5.1724
OP90	120	29.1594	4.1434	0.8756	0	0	0
OP100	113	28.1371	3.4827	0.8899	0	0	0
OP110	120	27.0904	3.4971	0.8857	0.3075	0.188	6.9386
OP120	105	99.9912	8.1551	0.9246	0	0	0

6.8.5 System Performance Analysis Using Four-weeks Averaged Data and Project Goal

The performance of the system at hand has been evaluated using the averaged data of Table 6.9 and the aggregation procedure of *Bai et al.* (2021). The result is shown in Figure 6.4. As one can see the bottleneck is OP10 and TP = 14.66JPH. Since the nominal throughput (defined by the longest cycle time, τ = 120sec, under the assumption that there are no machine breakdowns or cycle overruns) is 30JPH, the throughput losses are over 50%.

It is of interest to evaluate what fraction of these losses are due to machine downtime and due to cycle overrun. The former is calculated assuming that T_{down} of each machine is zero, and the latter that p_{OR} is zero. The throughput in the first case turns out to be 23.00JPH and in the second 17.26JPH. Thus, the production



	m_1	b_1	m_2	b_2	m_3	b_3	m_4	b_4	m_5	b_5	m_6	b_6	m_7	b_7	m_8	b_8	m_9	b_9	m_{10}	b_{10}	m_{11}	b_{11}	m_{12}
c (JPH)	30.00		30.25		30.00		30.00		33.96		30.00		30.00		30.00		30.00		31.86		30.00		34.29
e	0.74		0.75		0.73		0.84		0.90		0.87		0.93		0.84		0.88		0.89		0.89		0.92
T_{OR} (sec)	36.50		0.00		4.19		2.79		0.00		32.99		13.74		5.17		0.00		0.00		6.94		0.00
SAT (JPH)	16.91		22.67		21.09		24.53		30.46		20.23		24.93		24.14		26.27		28.35		25.11		31.70
WIP		0.67		1.07		0.72		0.22		0.87		0.63		0.29		0.18		0.27		0.35		0.18	
ST	0.00		0.21		0.18		0.27		0.40		0.22		0.30		0.29		0.37		0.41		0.36		0.50
BL	0.09		0.06		0.04		0.08		0.14		0.04		0.12		0.07		0.04		0.04		0.02		0.00
TP	14.6557JPH																						

Figure 6.4. Performance analysis based on the averaged data.

losses due to machine downtimes are 8.34JPH and due to cycle overrun 2.60JPH. Recovering these losses is the goal of this case study. More precisely, the goal is to design four options for a continuous improvement leading to a 5%, 10%, 20% and 30% of throughput increase.

6.8.6 Designing Continuous Improvement Projects

Applying the Continuous Improvement Design Procedure formulated in Section 6.6 to the production system at hand, we obtain incremental steps for continuous improvement resulting in up to 30% TP increase, i.e., TP_{imp} = 19.05JPH (see Table 6.10). Based on these incremental steps, we specify the activities to be carried out for each machine, leading to the desired throughput improvement. This results in the continuous improvement projects specified in Table 6.11. As one can see, for a 5% improvement, only one machine, OP10, must be improved; for 10% improvement, parameters of three machines, OP10, OP20, and OP60, should be modified; for 20% and 30% improvement, five machines, OP10, OP20, OP30, OP60, and OP80, should be improved (to various degrees).

This information is intended to allow the Operations Manager to decide which of these continuous improvement projects should be implemented on the factory floor.

Table 6.10. Incremental improvement steps.

	Improvement steps	TP	ΔTP	Percentage increase
0	initial state	14.6557	0	0%
1	reduce OP10 downtime by 10%	14.9317	0.276	1.88%
2	reduce OP10 downtime by 10%	15.1781	0.5224	3.56%
3	reduce OP10 cycle overrun by 20%	15.4996	0.8439	5.76%
4	reduce OP10 downtime by 10%	15.6961	1.0404	7.10%
5	reduce OP60 cycle overrun by 20%	15.796	1.1403	7.78%
6	reduce OP10 downtime by 10%	15.9698	1.3141	8.97%
7	reduce OP20 downtime by 10%	16.2481	1.5924	10.87%
8	reduce OP60 cycle overrun by 20%	16.3311	1.6754	11.43%
9	reduce OP10 cycle overrun by 20%	16.5204	1.8647	12.72%
10	reduce OP30 downtime by 10%	16.7583	2.1026	14.35%
11	reduce OP60 cycle overrun by 20%	16.826	2.1703	14.81%
12	reduce OP60 downtime by 10%	16.887	2.2313	15.22%
13	reduce OP20 downtime by 10%	17.1351	2.4794	16.92%
14	reduce OP30 downtime by 10%	17.3457	2.69	18.35%
15	reduce OP60 cycle overrun by 20%	17.3995	2.7438	18.72%
16	reduce OP60 downtime by 10%	17.4619	2.8062	19.15%

17	reduce OP80 downtime by 10%	17.5078	2.8521	19.46%
18	reduce OP10 downtime by 10%	17.6579	3.0022	20.49%
19	reduce OP30 downtime by 10%	17.8537	3.198	21.82%
20	reduce OP60 downtime by 10%	17.9111	3.2554	22.21%
21	reduce OP10 downtime by 10%	18.0441	3.3884	23.12%
22	reduce OP80 downtime by 10%	18.0959	3.4402	23.47%
23	reduce OP20 downtime by 10%	18.3165	3.6608	24.98%
24	reduce OP30 downtime by 10%	18.4906	3.8349	26.17%
25	reduce OP80 downtime by 10%	18.5484	3.8927	26.56%
26	reduce OP60 downtime by 10%	18.6059	3.9502	26.95%
27	reduce OP10 cycle overrun by 20%	18.7783	4.1226	28.13%
28	reduce OP30 downtime by 10%	18.9232	4.2675	29.12%
29	reduce OP80 downtime by 10%	18.9832	4.3275	29.53%
30	reduce OP60 cycle overrun by 20%	19.0356	4.3799	29.89%
31	reduce OP20 downtime by 10%	19.2177	4.562	31.13%

Table 6.11. Continuous improvement projects.

(a) For 5% throughput improvement (resulting, in fact, in 6%)

Machine	Machine improvement
OP10	Reduce downtime by 19%, reduce cycle overrun by 20%.

(b) For 10% throughput improvement, i.e., $TP_{imp} = 16.12$JPH

Machine	Machine improvement
OP10	Reduce downtime by 34%, reduce cycle overrun by 20%.
OP20	Reduce downtime by 10%.
OP60	Reduce cycle overrun by 20%.

(c) For 20% throughput improvement, i.e., $TP_{imp} = 17.59$JPH

Machine	Machine improvement
OP10	Reduce downtime by 41%, reduce cycle overrun by 36%.
OP20	Reduce downtime by 19%.
OP30	Reduce downtime by 19%.
OP60	Reduce downtime by 19%, reduce cycle overrun by 59%.
OP80	Reduce downtime by 10%.

(d) For 30% throughput improvement, i.e., $TP_{imp} = 19.05$JPH

Machine	Machine improvement
OP10	Reduce downtime by 47%, reduce cycle overrun by 49%.
OP20	Reduce downtime by 34%.
OP30	Reduce downtime by 41%.
OP60	Reduce downtime by 34%, reduce cycle overrun by 67%.
OP80	Reduce downtime by 34%.

6.9 Estimates of Cycle Overrun Parameters

6.9.1 Definition and Problem Formulation

Let $\widehat{p}_{OR}(n)$ be the random variable representing the estimate of the overrun probability p_{OR} based on n number of cycles measured:

$$\widehat{p}_{OR}(n) := \frac{n_{OR}}{n}, \tag{6.18}$$

where n_{OR} denotes the number of cycles where overrun takes place. Let $\widehat{k}_{OR}(n_{OR})$ be the estimate of k_{OR} based on measuring durations of n_{OR} number of cycle overrun (denoted as $t_{OR,i}$):

$$\widehat{k}_{OR}(n_{OR}) := \frac{\sum_{i=1}^{n_{OR}} t_{OR,i}}{n_{OR}\tau}. \tag{6.19}$$

Definition 6.2. We define the precision of the estimate $\widehat{p}_{OR}(n)$ and $\widehat{k}_{OR}(n_{OR})$ as follows:

- The estimate $\widehat{p}_{OR}(n)$ is referred to as (α, β)-precise if

$$P\left\{\frac{|\widehat{p}_{OR}(n) - p_{OR}|}{p_{OR}} \leqslant \alpha\right\} \geqslant \beta. \tag{6.20}$$

- The estimate $\widehat{k}_{OR}(n_{OR})$ is referred to as (α, β)-precise if

$$P\left\{\frac{|\widehat{k}_{OR}(n_{OR}) - k_{OR}|}{k_{OR}} \leqslant \alpha\right\} \geqslant \beta. \tag{6.21}$$

Definition 6.3. The smallest integers $n^*_{p_{OR}}(\alpha, \beta)$ and $n^*_{k_{OR}}(\alpha, \beta)$, which guarantees

(6.20) and (6.21), respectively, are referred to as the critical numbers of cycle overrun measurements for p_{OR} and k_{OR}, respectively.

Consider the unconditional cycle overrun, the estimate of its mean is calculated as

$$\widehat{T}_{OR} = \widehat{p}_{OR}\widehat{k}_{OR}\tau, \tag{6.22}$$

and its precision is characterized as

$$P\left\{\frac{|\widehat{T}_{OR} - T_{OR}|}{T_{OR}} \leqslant \alpha_{OR}\right\} \geqslant \beta_{OR}. \tag{6.23}$$

We introduce the following problems:

Induced Precision Problem for $\boldsymbol{T_{OR}}$: Given (α_1, β_1)-precise estimate $\widehat{p}_{OR}\left(n^*_{p_{OR}}(\alpha_1, \beta_1)\right)$ and (α_2, β_2)-precise estimate $\widehat{k}_{OR}\left(n^*_{k_{OR}}(\alpha_2, \beta_2)\right)$, calculate the induced precision $(\alpha^{ind}_{OR}, \beta^{ind}_{OR})$ of $\widehat{T}_{OR}$.

Inverse Problem for $\boldsymbol{T_{OR}}$: Given a desired (γ, δ), calculate the smallest number of cycle overrun measurements, $n^{**}_{OR}(\gamma, \delta)$, required to obtain (γ, δ)-precise estimate of T_{OR}.

6.9.2 Critical Number of Cycle Overrun Measurements

Proposition 6.4. *The Gaussian approximation of critical number $n^*_{p_{OR}}(\alpha, \beta)$ is given by*

$$n^*_{p_{OR}}(\alpha, \beta) = \left\lceil 2\left(\frac{1 - \widehat{p}_{OR}}{\widehat{p}_{OR}}\right)\left(\frac{\operatorname{erf}^{-1}(\beta)}{\alpha}\right)\right\rceil. \tag{6.24}$$

Justification. Similar to the justification of Proposition 4.1. □

Proposition 6.5. *The Gaussian approximation of critical number* $n^*_{k_{OR}}(\alpha, \beta)$ *is given by*

$$n^*_{k_{OR}}(\alpha, \beta) = \left\lceil 2 \left(\frac{\text{erf}^{-1}(\beta)}{\alpha} \right)^2 \right\rceil. \tag{6.25}$$

Proof. See Appendix E. □

6.9.3 Induced Precision and Inverse Problem for T_{OR}

The solution to the induced precision problem is given by Propositions 6.6 and 6.7:

Proposition 6.6. *Given* $\frac{|\hat{p}_{OR}(n^*_{p_{OR}}) - p_{OR}|}{p_{OR}} \leqslant \alpha_1$ *and* $\frac{|\hat{k}_{OR}(n^*_{k_{OR}}) - k_{OR}|}{k_{OR}} \leqslant \alpha_2$*, the smallest induced* α_{OR} *which satisfies* $\frac{|\hat{T}_{OR} - T_{OR}|}{T_{OR}} \leqslant \alpha^{ind}_{OR}$ *with accuracy* $O(\alpha_1\alpha_2)$ *is given by*

$$\alpha^{ind}_{OR} = \alpha_1 + \alpha_2. \tag{6.26}$$

Justification. See Appendix E. □

Proposition 6.7. *The Gaussian approximation of* β^{ind}_{OR} *is given by*

$$\beta^{ind}_{OR} = \text{erf}\left((\alpha_1 + \alpha_2) C \sqrt{\frac{n^*_{p_{OR}}}{2}} \right), \tag{6.27}$$

where $C = \sqrt{\frac{\hat{p}_{OR}\left(n^*_{p_{OR}}\right)}{2 - \hat{p}_{OR}\left(n^*_{p_{OR}}\right)}}$.

Justification. See Appendix E. □

The solution to the inverse problem is given by Proposition 6.8:

Proposition 6.8. *For a given (γ, δ), the critical number of cycle overrun measurements to ensure (γ, δ)-precise estimate of T_{OR} is given by*

$$n_{OR}^{**}(\gamma, \delta) = \left\lceil 2 \left(\frac{2 - \widehat{p}_{OR}}{\widehat{p}_{OR}} \right) \left(\frac{\text{erf}^{-1}(\delta)}{\gamma} \right) \right\rceil, \tag{6.28}$$

where $\widehat{p}_{OR} = \widehat{p}_{OR}\left(n_{OR}^{**}(\gamma, \delta)\right)$.

Justification. See Appendix E. □

Comparing (6.28) with critical numbers given by (6.24) and (6.25), we conclude:

Observation 6.1. The performance metric T_{OR} is variability expanding.

As one can see, (6.28) is not in closed-form, thus, the Approximation Procedure of Chapter 3 should be used. Here we consider using Approximation Procedure 1. After experimentation, we find that the initial number of cycle overrun measurements $n_{OR,ini}$ and the safety factor ϵ can be selected as 30 and 0.05. The following numerical example shows the effectiveness of the Approximation Procedure:

We generate 25 machines with cycle overrun, with parameters randomly and equiprobabily selected from the following sets:

$$\tau \in [1, 2], \; p_{OR} \in [0.3, 1], \; T_{OR}^c = k_{OR}\tau, \text{ where } k = [0.2, 2].$$

The desired precision pair (γ, δ) is selected as $(0.1, 0.9)$, $(0.1, 0.95)$, $(0.05, 0.9)$, and $(0.05, 0.95)$. The parameters of the five randomly generated 25 machines are listed in Table 6.12. For the five machines and precision pairs (γ, δ), the values of $\widehat{\delta}$ are listed in Table 6.13, the corresponding values of n_{OR}^{**} and $\breve{n}_{OR}^{**}\left(\gamma, \delta; \widehat{p}_{OR}(n_{ini}), \epsilon\right)$ are listed in Table 6.14. As we can see, for the five randomly generated machines, the precision of the estimate $\widehat{T}_{OR}\left(\breve{n}^{**}_{OR}\left(\gamma,\delta;\widehat{p}_{OR}(n_{ini}),\epsilon\right)\right)$ is always higher than the desired (γ,δ), and the values of $\breve{n}^{**}_{OR}\left(\gamma,\delta;\widehat{p}_{OR}(n_{ini}),\epsilon\right)$ are only about 10% to 15% larger than its true values. Similar results are obtained for the rest 20 machines as well.

Table 6.12. Parameters of five randomly generated machines.

Parameter \ Machine	$\boldsymbol{m_1}$	$\boldsymbol{m_2}$	$\boldsymbol{m_3}$	$\boldsymbol{m_4}$	$\boldsymbol{m_5}$
τ	1.7636	1.0458	1.8439	1.8552	1.2197
p_{OR}	0.4609	0.7872	0.7535	0.9640	0.6839
T_{OR}	2.3597	1.0199	3.6003	0.6100	0.9119

Table 6.13. Values of $\widehat{\delta}$ as a function of (γ,δ).

$(\boldsymbol{\gamma},\boldsymbol{\delta})$ \ Machine	$\boldsymbol{m_1}$	$\boldsymbol{m_2}$	$\boldsymbol{m_3}$	$\boldsymbol{m_4}$	$\boldsymbol{m_5}$
(0.1, 0.9)	0.9083	0.9104	0.9188	0.9138	0.9117
(0.1, 0.95)	0.9529	0.9566	0.9554	0.9605	0.9563
(0.05, 0.9)	0.9079	0.9164	0.9096	0.9147	0.9142
(0.05, 0.95)	0.9523	0.9600	0.9607	0.9568	0.9550

Table 6.14. Values of $n^{**}_{OR}(\gamma,\delta)$ and $\breve{n}^{**}_{OR}\left(\gamma,\delta;\widehat{p}_{OR}(n_{ini}),\epsilon\right)$ as functions of (γ,δ).

Machine	$\boldsymbol{m_1}$		$\boldsymbol{m_2}$		$\boldsymbol{m_3}$		$\boldsymbol{m_4}$		$\boldsymbol{m_5}$	
$(\boldsymbol{\gamma},\boldsymbol{\delta})$	n^{**}_{OR}	$\breve{n}^{**}_{OR}$	n^{**}_{OR}	$\breve{n}^{**}_{OR}$	n^{**}_{OR}	$\breve{n}^{**}_{OR}$	n^{**}_{OR}	$\breve{n}^{**}_{OR}$	n^{**}_{OR}	$\breve{n}^{**}_{OR}$
(0.1, 0.9)	904	1021	417	462	448	494	291	322	521	577
(0.1, 0.95)	1283	1446	592	654	636	702	413	457	740	818
(0.05, 0.9)	3614	4085	1668	1838	1791	1975	1164	1286	2083	2304
(0.05, 0.95)	5131	5770	2368	2618	2543	2803	1652	1825	2957	3277

6.10 Summary

This chapter provides analytical methods for analysis and improvement of serial lines with unreliable machines and cycle overrun. These methods offer the analytics for modelling, analysis and improvement for a relatively large class of real-world serial lines, which has not been thus far explored in production systems literature. In addition, the (α, β)-precision theory is applied to obtain the critical numbers of cycle overrun measurements ($n^*_{p_{OR}}$ and $n^*_{k_{OR}}$) and the solutions to the induced precision and inverse problems for T_{OR}.

Results reported here can be extended in at least two directions. The first one is to develop similar methods for a wider class of systems, namely, serial lines with rework, with carriers, and with quality issues, as well as for assembly systems. The second is to further explore production systems with non-exponential machines and non-exponential cycle overrun. Several results in this direction are included in Section 6.7, leading to the hypotheses formulated in Subsections 6.7.2 and 6.7.3. Proving these hypotheses would substantially extend the "safety" of using the exponential assumption in designing and improving production systems. In addition to these two areas of future research, a very important one is the application of the methods developed to production systems on the factory floor.

CHAPTER 7

Conclusion and Future Work

7.1 Conclusion

The contributions of this dissertation include:

- A method for calculating the $n_T^*(\alpha, \beta)$ of up- and downtime measurements, necessary and sufficient to obtain (α, β)-precise estimates of $MTBF$ and $MTTR$. Extended the results to non-exponential case with $CV < 1$.

- A method for evaluating the observation time required to collect $n_T^*(\alpha, \beta)$ measurements of machine up- and downtime.

- Evaluation of the induced precision of machine efficiency (e), throughput (TP), production lead time (LT), and work-in-process (WIP) estimates.

- Solution to the inverse problem for machine efficiency, throughput, production lead time, and work-in-process.

- A method for calculating the $n_q^*(\alpha_q, \beta_q)$ of quality parts measurements, required to obtain (α_q, β_q)-precise estimate of machine quality parameter (q).

- Evaluation of the induced precision of quality parts throughput (TP_q) estimate, solution to the inverse problem concerning TP_q.

- Analysis of variability contracting/expanding property for e, TP, LT, WIP, and TP_q.

- Modeling, analysis, bottleneck identification and improvement of production systems with cycle overrun.

7.2 Future Work

The work in this dissertation can be extended in the following directions:

- Extend the analysis and evaluation of the critical numbers to assembly system performance metrics, and to performance metrics of production systems with non-exponential machines.

- Extend the analysis, improvability and bottleneck identification to a wider class of systems with cycle overrun, namely, serial lines with quality issues and assembly systems.

- Apply the (α, β)-Precision Theory results to practice and merge them into automated production system monitoring and improvement software.

APPENDICES

APPENDIX A

Appendix of Chapter 2

A.1 Proof of Theorem 2.1

First, we prove the statement of this theorem for $\widehat{T}_{up}$, and then address $\widehat{T}_{down}$.

Note that $t_{up,i}$'s are iid exponential random variables with parameter $\lambda = 1/T_{up}$. This implies that $\widehat{T}_{up}(n)$ is an Erlang random variable with shape parameter n and rate parameter $n\lambda$. Its cdf is given by

$$F_{\widehat{T}_{up}(n)}(x) = P\{\widehat{T}_{up}(n) \leqslant x\} = 1 - \sum_{i=0}^{n-1} \frac{1}{i!} e^{-n\lambda x}(n\lambda x)^i. \tag{A.1}$$

Then, from the first expression in (2.3), we obtain:

$$
\begin{aligned}
& P\left\{(1-\alpha)T_{up} \leqslant \widehat{T}_{up}(n) \leqslant (1+\alpha)T_{up}\right\} \\
=& F_{\widehat{T}_{up}(n)}\left((1+\alpha)T_{up}\right) - F_{\widehat{T}_{up}(n)}\left((1-\alpha)T_{up}\right) \\
=& \left(1 - \sum_{i=0}^{n-1} \frac{1}{i!} e^{-n\lambda(1+\alpha)T_{up}} \left(n\lambda(1+\alpha)T_{up}\right)^i\right) \\
& - \left(1 - \sum_{i=0}^{n-1} \frac{1}{i!} e^{-n\lambda(1-\alpha)T_{up}} \left(n\lambda(1-\alpha)T_{up}\right)^i\right) \\
=& \sum_{i=0}^{n-1} \frac{1}{i!} e^{-(1-\alpha)n}((1-\alpha)n)^i - \sum_{i=0}^{n-1} \frac{1}{i!} e^{-(1+\alpha)n}((1+\alpha)n)^i \geqslant \beta.
\end{aligned}
\tag{A.2}
$$

Therefore, the critical number $n^*(\alpha, \beta)$ is the smallest integer n, which satisfies the inequality:

$$
\beta \leqslant \sum_{i=0}^{n-1} \frac{1}{i!} e^{-(1-\alpha)n}((1-\alpha)n)^i - \sum_{i=0}^{n-1} \frac{1}{i!} e^{-(1+\alpha)n}((1+\alpha)n)^i. \tag{A.3}
$$

Since this expression is independent of λ, it holds also for μ, i.e., for T_{down}.

A.2 Justification of Proposition 2.3

According to the central limit theorem, for large $n^*(\alpha, \beta)$, Erlang random variable $\widehat{T}_{up}(n_T^*)$ can be approximated by the Gaussian distribution, with mean $M = \frac{1}{\lambda} = T_{up}$ and variance $V = \frac{1}{\lambda^2 n^*} = \frac{T_{up}^2}{n^*}$. Therefore, the probability in (2.3) can be approximated as:

$$\begin{aligned}
& P\left\{(1-\alpha)T_{up} \leqslant \widehat{T}_{up}(n_T^*) \leqslant (1+\alpha)T_{up}\right\} \\
= \quad & P\left\{\frac{(1-\alpha)T_{up}-T_{up}}{\frac{T_{up}}{\sqrt{n^*}}} \leqslant \frac{\widehat{T}_{up}(n_T^*)-T_{up}}{\frac{T_{up}}{\sqrt{n^*}}} \leqslant \frac{(1+\alpha)T_{up}-T_{up}}{\frac{T_{up}}{\sqrt{n^*}}}\right\} \qquad \text{(A.4)} \\
\approx \quad & P\left\{|Z| \leqslant \alpha\sqrt{n^*}\right\},
\end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$ denotes the standard normal distribution. Therefore, the Gaussian approximation of $n^*(\alpha,\beta)$ can be obtained from the following equation:

$$\beta = \int_{-\alpha\sqrt{n_G^*}}^{\alpha\sqrt{n_G^*}} f_Z(z)dz = \operatorname{erf}\left(\frac{\alpha\sqrt{n_G^*}}{\sqrt{2}}\right), \qquad \text{(A.5)}$$

i.e.,

$$n_G^*(\alpha,\beta) = \left\lceil 2\left(\frac{\operatorname{erf}^{-1}(\beta)}{\alpha}\right)^2\right\rceil. \qquad \text{(A.6)}$$

□

APPENDIX B

Appendix of Chapter 3

B.1 Justification of Proposition 3.1

Denote $\rho = \frac{T_{down}}{T_{up}}$ and $\widehat{\rho} = \frac{\widehat{T}_{down}}{\widehat{T}_{up}}$. Since $(1-\alpha)T_{up} \leqslant \widehat{T}_{up} \leqslant (1+\alpha)T_{up}$, and $0 < (1-\alpha)T_{down} \leqslant \widehat{T}_{down} \leqslant (1+\alpha)T_{down}$, we have:

$$\begin{aligned}
&\tfrac{(1-\alpha)T_{up}}{(1+\alpha)T_{down}} \leqslant \tfrac{\widehat{T}_{up}}{\widehat{T}_{down}} \leqslant \tfrac{(1+\alpha)T_{up}}{(1-\alpha)T_{down}} \\
&\Leftrightarrow \tfrac{1-\alpha}{1+\alpha}\tfrac{1}{\rho} \leqslant \tfrac{1}{\widehat{\rho}} \leqslant \tfrac{1+\alpha}{1-\alpha}\tfrac{1}{\rho} \\
&\Leftrightarrow \tfrac{1-\alpha}{1+\alpha}\widehat{\rho} \leqslant \rho \leqslant \tfrac{1+\alpha}{1-\alpha}\widehat{\rho} \\
&\Leftrightarrow \tfrac{-2\alpha}{1+\alpha}\widehat{\rho} \leqslant \rho - \widehat{\rho} \leqslant \tfrac{2\alpha}{1-\alpha}\widehat{\rho}.
\end{aligned} \tag{B.1}$$

Dividing (B.1) by $(1+\rho)(1+\widehat{\rho})$, and substituting $\frac{1}{1+\rho}$ with e, and $\frac{1}{1+\widehat{\rho}}$ with $\widehat{e}$, we obtain:

$$\frac{-2\alpha}{1+\alpha}e(1-\widehat{e}) \leqslant \widehat{e} - e \leqslant \frac{2\alpha}{1-\alpha}e(1-\widehat{e}). \tag{B.2}$$

For small α, using Taylor expansion we have:

$$-2\alpha e(1-\widehat{e}) + O(\alpha^2) \leqslant \widehat{e} - e \leqslant 2\alpha e(1-\widehat{e}) + O(\alpha^2). \tag{B.3}$$

Therefore, we obtain:

$$\frac{|e-\widehat{e}|}{e} \leqslant \alpha_e = 2(1-\widehat{e})\alpha + O(\alpha^2). \tag{B.4}$$

□

B.2 Proof of Theorem 3.2

As discussed in Appendix A, $Y(n) = \sum_{i=1}^{n} t_{up,i}$ obeys Erlang distribution with shape parameter n and scale parameter T_{up}. Therefore, the pdf of $Y(n)$ is

$$f_{Y(n)}(x) = (\frac{1}{T_{up}})^n \cdot \frac{x^{n-1}e^{-\frac{x}{T_{up}}}}{(n-1)!},$$

and the pdf of $\widehat{T}_{up}(n) = \frac{Y(n)}{n}$ is

$$f_{\widehat{T}_{up}(n)}(x) = n \cdot (\frac{1}{T_{up}})^n \cdot \frac{(n \cdot x)^{n-1}e^{-\frac{n \cdot x}{T_{up}}}}{(n-1)!}.$$

Similarly,

$$f_{\widehat{T}_{down}(n)}(x) = n \cdot (\frac{1}{T_{down}})^n \cdot \frac{(n \cdot x)^{n-1}e^{-\frac{n \cdot x}{T_{down}}}}{(n-1)!}.$$

Since $\widehat{T}_{up}(n)$ and $\widehat{T}_{down}(n)$ are independent,

$$f_{\widehat{T}_{up}(n),\widehat{T}_{down}(n)}(x,y) = f_{\widehat{T}_{up}(n)}(x) \cdot f_{\widehat{T}_{down}(n)}(y).$$

Using the notations $\rho = \frac{T_{down}}{T_{up}}$ and $\widehat{\rho} = \frac{\widehat{T}_{down}(n)}{\widehat{T}_{up}(n)}$, we can write:

$$\begin{aligned} & P\left\{\tfrac{|e-\hat{e}(n^*)|}{e} \leqslant \alpha_e\right\} \\ = \; & P\left\{-2\alpha\big(1-\widehat{e}(n^*)\big)e \leqslant e - \widehat{e}(n^*) \leqslant 2\alpha\big(1-\widehat{e}(n^*)\big)e\right\} \\ = \; & P\left\{2\alpha\tfrac{\widehat{\rho}}{1+\widehat{\rho}}\tfrac{1}{1+\rho} \leqslant \tfrac{1}{1+\rho} - \tfrac{1}{1+\widehat{\rho}} \leqslant 2\alpha\tfrac{\widehat{\rho}}{1+\widehat{\rho}}\tfrac{1}{1+\rho}\right\} \\ = \; & P\left\{\tfrac{1}{1+2\alpha}\rho \leqslant \widehat{\rho} \leqslant \tfrac{1}{1-2\alpha}\rho\right\}. \end{aligned} \tag{B.5}$$

Denote $\rho_1 = \frac{1}{1+2\alpha}\rho$, and $\rho_2 = \frac{1}{1-2\alpha}\rho$, we derive:

$$\begin{aligned} & P\left\{\tfrac{|e-\hat{e}(n^*)|}{e} \leqslant \alpha_e\right\} \\ = \; & \textstyle\int_0^{+\infty}\int_{\rho_1 x}^{\rho_2 x} f_{\widehat{T}_{up}(n^*)}(x) f_{\widehat{T}_{down}(n^*)}(y)dydx \\ = \; & \left(\tfrac{{n^*}^2}{T_{up}T_{down}}\right)^{n^*}\left(\tfrac{1}{(n^*-1)!}\right)^2 \textstyle\int_0^{+\infty} x^{n^*-1}e^{-\frac{n^*}{T_{up}}x}\int_{\rho_1 x}^{\rho_2 x} y^{n^*-1}e^{-\frac{n^*}{T_{down}}y}dydx, \end{aligned} \tag{B.6}$$

where

$$\begin{aligned} & \textstyle\int_{\rho_1 x}^{\rho_2 x} y^{n^*-1}e^{-\frac{n^*}{T_{down}}y}dy \\ = \; & e^{-\frac{n^*}{T_{down}}y}\textstyle\sum_{i=0}^{n^*-1}\tfrac{(n^*-1)!}{(n^*-1-i)!}\left(\tfrac{T_{down}}{n^*}\right)^{i+1}y^{n^*-1-i}\Big|_{\rho_1 x}^{\rho_2 x} \\ = \; & e^{-\frac{n^*}{T_{down}}\rho_1 x}\textstyle\sum_{i=0}^{n^*-1}\tfrac{(n^*-1)!}{(n^*-1-i)!}\left(\tfrac{T_{down}}{n^*}\right)^{i+1}(\rho_1 x)^{n^*-1-i} - \\ & e^{-\frac{n^*}{T_{down}}\rho_2 x}\textstyle\sum_{i=0}^{n^*-1}\tfrac{(n^*-1)!}{(n^*-1-i)!}\left(\tfrac{T_{down}}{n^*}\right)^{i+1}(\rho_2 x)^{n^*-1-i}. \end{aligned} \tag{B.7}$$

Therefore,

$$
\begin{aligned}
& P\left\{\tfrac{|e-\widehat{e}(n^*)|}{e} \leqslant \alpha_e\right\} \\
= \quad & (\tfrac{{n^*}^2}{T_{up}T_{down}})^{n^*}(\tfrac{1}{(n^*-1)!})^2 \int_0^{+\infty} x^{n^*-1} e^{-\frac{n^*}{T_{up}}x} \\
& \Big(e^{-\frac{n^*}{T_{down}}\rho_1 x} \textstyle\sum_{i=0}^{n^*-1} \tfrac{(n^*-1)!}{(n^*-1-i)!}(\tfrac{T_{down}}{n^*})^{i+1}(\rho_1 x)^{n^*-1-i} - \\
& e^{-\frac{n^*}{T_{down}}\rho_2 x} \textstyle\sum_{i=0}^{n^*-1} \tfrac{(n^*-1)!}{(n^*-1-i)!}(\tfrac{T_{down}}{n^*})^{i+1}(\rho_2 x)^{n^*-1-i}\Big) dx.
\end{aligned}
\tag{B.8}
$$

Denote

$$
\begin{aligned}
\mathcal{A} = \quad & \textstyle\int_0^{+\infty} x^{n^*-1} e^{-\frac{n^*}{T_{up}}x} \Big(e^{-\frac{n^*}{T_{down}}\rho_1 x} \sum_{i=0}^{n^*-1} \tfrac{(n^*-1)!}{(n^*-1-i)!} \times \\
& (\tfrac{T_{down}}{n^*})^{i+1}(\rho_1 x)^{n^*-1-i}\Big) dx,
\end{aligned}
\tag{B.9}
$$

and

$$
\begin{aligned}
\mathcal{B} = \quad & \textstyle\int_0^{+\infty} x^{n^*-1} e^{-\frac{n^*}{T_{up}}x} \Big(e^{-\frac{n^*}{T_{down}}\rho_2 x} \sum_{i=0}^{n^*-1} \tfrac{(n^*-1)!}{(n^*-1-i)!} \times \\
& (\tfrac{T_{down}}{n^*})^{i+1}(\rho_2 x)^{n^*-1-i}\Big) dx,
\end{aligned}
\tag{B.10}
$$

we can write:

$$
P\left\{\frac{|e-\widehat{e}(n^*)|}{e} \leqslant \alpha_e\right\} = (\frac{{n^*}^2}{T_{up}T_{down}})^{n^*}(\frac{1}{(n^*-1)!})^2\big(\mathcal{A}-\mathcal{B}\big). \tag{B.11}
$$

We obtain:

$$\begin{aligned}\mathcal{A} =& \sum_{i=0}^{n^*-1} \frac{(n^*-1)!}{(n^*-1-i)!}(\frac{T_{down}}{n^*})^{i+1}\rho_1^{n^*-1-i}\int_0^{+\infty} e^{-(\frac{n^*}{T_{up}}+\frac{n^*}{T_{down}}\rho_1)x}x^{2n^*-2-i}dx \\ =& \sum_{i=0}^{n^*-1} \frac{(n^*-1)!}{(n^*-1-i)!}(\frac{T_{down}}{n^*})^{i+1}\rho_1^{n^*-1-i}\Bigg(e^{-(\frac{n^*}{T_{up}}+\frac{n^*}{T_{down}}\rho_1)x}\times \\ & \sum_{j=0}^{a}(-1)^j \frac{a!}{(a-j)!\left[-(\frac{n^*}{T_{up}}+\frac{n^*}{T_{down}}\rho_1)\right]^{j+1}}x^{a-j}|_0^{+\infty}\Bigg) \\ =& n^{*-2n^*}\sum_{i=0}^{n^*-1}\frac{(n^*-1)!(2n^*-2-i)!}{(n^*-1-i)!}T_{up}^{n^*}T_{down}^{n^*}(1+2\alpha)^{n^*}(2+2\alpha)^{-2n^*+i+1}.\end{aligned} \tag{B.12}$$

Similarly,

$$\mathcal{B} = \quad n^{*-2n^*}\sum_{i=0}^{n^*-1}\frac{(n^*-1)!(2n^*-2-i)!}{(n^*-1-i)!}T_{up}^{n^*}T_{down}^{n^*}(1-2\alpha)^{n^*}(2-2\alpha)^{-2n^*+i+1}. \tag{B.13}$$

Substitute values of $\mathcal{A}$ and $\mathcal{B}$ in (B.11), we get:

$$\begin{aligned}& P\left\{\frac{|e-\widehat{e}(n^*)|}{e} \leqslant \alpha_e\right\} \\ =& \sum_{i=0}^{n^*-1}\frac{(2n^*-2-i)!}{(n^*-1-i)!(n^*-1)!}\Big[(1+2\alpha)^{n^*}(2+2\alpha)^{-2n^*+i+1} \\ & -(1-2\alpha)^{n^*}(2-2\alpha)^{-2n^*+i+1}\Big].\end{aligned} \tag{B.14}$$

Thus, by (3.2), we have:

$$\begin{aligned}\beta =& \sum_{i=0}^{n^*-1}\frac{(2n^*-2-i)!}{(n^*-1-i)!(n^*-1)!}\Big[(1+2\alpha)^{n^*}(2+2\alpha)^{-2n^*+i+1} \\ & -(1-2\alpha)^{n^*}(2-2\alpha)^{-2n^*+i+1}\Big].\end{aligned} \tag{B.15}$$

□

B.3 Justification of Proposition 3.3

Consider the estimate of $\widehat{e}(n_T^*)$ given by:

$$\widehat{e}(n_T^*) = \frac{\widehat{T}_{up}(n_T^*)}{\widehat{T}_{up}(n_T^*) + \widehat{T}_{down}(n_T^*)}, \tag{B.16}$$

where $\widehat{T}_{up}$ and $\widehat{T}_{down}$ are (α, β)-precise estimates of T_{up} and T_{down}, and n_T^* is the critical number corresponding to (α, β). From the pdfs of $\widehat{T}_{up}$ and $\widehat{T}_{down}$ given by

$$\begin{aligned} f_{\widehat{T}_{up}(n_T^*)}(x) &= n_T^* \cdot \left(\frac{1}{T_{up}}\right)^{n_T^*} \cdot \frac{(n_T^* \cdot x)^{n_T^*-1} e^{-\frac{n_T^* \cdot x}{T_{up}}}}{(n_T^*-1)!}, \\ f_{\widehat{T}_{down}(n_T^*)}(x) &= n_T^* \cdot \left(\frac{1}{T_{down}}\right)^{n_T^*} \cdot \frac{(n_T^* \cdot x)^{n_T^*-1} e^{-\frac{n_T^* \cdot x}{T_{down}}}}{(n_T^*-1)!}, \end{aligned} \tag{B.17}$$

we can derive the pdf of $\widehat{e}(n_T^*)$:

$$f_{\widehat{e}(n_T^*)}(x) = \frac{(2n_T^*-1)!}{(n_T^*-1)!^2} \frac{1}{x(1-x)} \left(\frac{T_{up}}{T_{down}} \left(\frac{1}{x} - 1\right) + 2 + \frac{T_{down}}{T_{up}} \left(\frac{1}{x} - 1\right)^{-1} \right)^{-n_T^*}, x \in (0,1). \tag{B.18}$$

Being non-analytic, this expression is not convenient for analysis and calculations. Therefore, we use two approximations of (B.16), leading to analytic pdf's. The first one is based on representing (B.16) as

$$\hat{e}_{NL}(n_T^*) = \frac{\widehat{T}_{up,G}(n_T^*)}{\widehat{T}_{up,G}(n_T^*) + \widehat{T}_{down,G}(n_T^*)}, \tag{B.19}$$

where $\widehat{T}_{up,G}(n_T^*)$ and $\widehat{T}_{down,G}(n_T^*)$ are Gaussian random variables with the pdf's $\mathcal{N}(T_{up}, \frac{T_{up}^2}{n_T^*})$ and $\mathcal{N}(T_{down}, \frac{T_{down}^2}{n_T^*})$, respectively, and NL stands for "nonlinear", indicating that (B.19) is a nonlinear function of its arguments. Using the formula for pdf of ratio of Gaussian random variables (*Marsaglia* 2006), we obtain:

$$f_{\hat{e}_{NL}(n_T^*)}(x) = \frac{T_{up}^2+T_{down}^2}{T_{up}\cdot T_{down}} f_W\left(\frac{T_{up}}{T_{down}} - \frac{T_{up}\cdot T_{down}}{T_{up}^2+T_{down}^2}\cdot x\right), \tag{B.20}$$

where

$$\begin{aligned} f_W(w) = \; & \frac{\exp(-n_T^*)}{(w^2+1)\pi}\Bigg\{1 + \sqrt{\frac{\pi}{2(1+w^2)}}\operatorname{erf}\Bigg(\sqrt{\frac{n_T^*}{2(1+w^2)(T_{up}^2+T_{down}^2)}} \\ & \cdot\Big(T_{up} + T_{down} + \big(T_{up} - T_{down}\big)w\Big)\Bigg)\exp\left(\frac{n_T^*\Big((T_{up}+T_{down})+(T_{up}+T_{down})w\Big)^2}{2(T_{up}+T_{down})(1+w^2)}\right) \\ & \cdot\left(\frac{\sqrt{n_T^*}(T_{up}+T_{down})}{\sqrt{T_{up}^2+T_{down}^2}} + \frac{\sqrt{n_T^*}(T_{up}-T_{down})}{\sqrt{T_{up}^2+T_{down}^2}}w\right)\Bigg\}. \end{aligned} \tag{B.21}$$

While this pdf is indeed analytic, it is relatively complex. Therefore, we use Taylor expansion of (B.19) in order to obtain its linear approximation. As a result, with the accuracy up to $O((\widehat{T}_{up,G}(n_T^*) - T_{up})^2)$ and $O((\widehat{T}_{down,G}(n_T^*) - T_{down})^2)$, we obtain:

$$\widehat{e}_L(n_T^*) = e + \tfrac{T_{down}}{(T_{up}+T_{down})^2} \left(\widehat{T}_{up,G}(n_T^*) - T_{up}\right) + \tfrac{-T_{up}}{(T_{up}+T_{down})^2} \left(\widehat{T}_{down,G}(n_T^*) - T_{down}\right), \tag{B.22}$$

where L stands for "linear". Since $\widehat{T}_{up,G}(n_T^*)$ and $\widehat{T}_{down,G}(n_T^*)$ are Gaussian, $\widehat{e}_L(n_T^*)$ is also Gaussian with mean and variance given by

$$E[\widehat{e}_L(n_T^*)] = e, \qquad Var[\widehat{e}_L(n_T^*)] = \frac{2}{n_T^*} e^2(1-e)^2. \tag{B.23}$$

In the subsequent discussions, we denote $\widehat{e}_L(n_T^*)$ as $\widehat{e}_G(n_T^*)$. Its pdf and cdf are given by:

$$f_{\widehat{e}_G(n_T^*)}(x) = \frac{\sqrt{n_T^*}}{2\sqrt{\pi}e(1-e)} \exp\left(-\frac{n_T^*(x-e)^2}{4e^2(1-e)^2}\right), \tag{B.24}$$

$$F_{\widehat{e}_G(n_T^*)}(x) = \frac{1}{2}\left(1 + \text{erf}\left(\frac{\sqrt{n_T^*}(x-e)}{2e(1-e)}\right)\right). \tag{B.25}$$

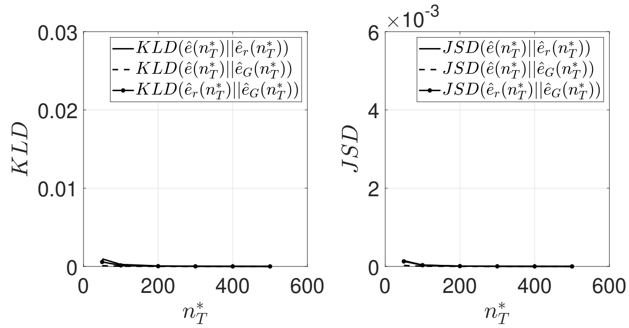
To select which one of the two approximations, (B.20), (B.21) or (B.24), should be used, we investigate the "distance" of both from the original one, (B.18). This investigation is carried out using the Kullback-Leibler divergence (KLD) and its symmetrized version Jensen-Shannon divergence (JSD) (*Kullback* 1997; *Lin* 1991) defined by:

$$\begin{aligned} KLD(X||Y) &:= \textstyle\int_{-\infty}^{\infty} f_X(z) log \left(\frac{f_X(z)}{f_Y(z)}\right) dz, \\ JSD(X||Y) &:= \tfrac{1}{2} KLD\left(X||\tfrac{X+Y}{2}\right) + \tfrac{1}{2} KLD\left(Y||\tfrac{X+Y}{2}\right), \end{aligned} \tag{B.26}$$

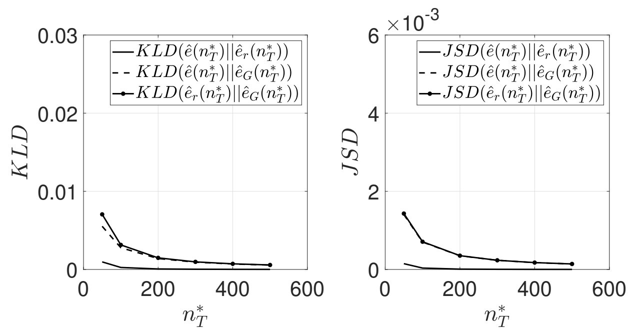
where X and Y are random variables with pdfs $f_X(z)$ and $f_Y(z)$, respectively.

To quantify the relationships within all pairs of random variables comprised of $\widehat{e}_{NL}(n_T^*)$, $\widehat{e}_G(n_T^*)$ and $\widehat{e}(n_T^*)$, we calculate KLD and JSD numerically for $n_T^* = \{50, 100, 200, 300, 400, 500\}$ and $e = \{0.5, 0.7, 0.9\}$. The results are shown in Figure B.1. As one can see:

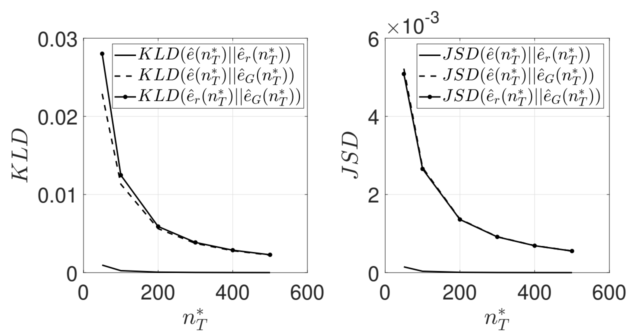
- $JSD(\widehat{e}(n_T^*)||\widehat{e}_{NL}(n_T^*))$ is smaller than $JSD(\widehat{e}(n_T^*)||\widehat{e}_G(n_T^*))$;
- $JSD(\widehat{e}(n_T^*)||\widehat{e}_G(n_T^*))$ is approximately equal to $JSD(\widehat{e}_{NL}(n_T^*)||\widehat{e}_G(n_T^*))$;
- however, for all pairs $(\widehat{e}(n_T^*), \widehat{e}_{NL}(n_T^*))$, $(\widehat{e}(n_T^*), \widehat{e}_G(n_T^*))$ and $(\widehat{e}_{NL}(n_T^*), \widehat{e}_G(n_T^*))$, the values of JSD are quite small; indeed, while maximum of $JSD = \ln 2$, the values of JSD for all three pairs are three orders of magnitude smaller.



(a) $e = 0.5$



(b) $e = 0.7$



(c) $e = 0.9$

Figure B.1. KLD and JSD between $\widehat{e}(n_T^*)$, $\widehat{e}_{NL}(n_T^*)$ and $\widehat{e}_G(n_T^*)$

Thus, from the point of view of JSD, either $\widehat{e}_G(n_T^*)$ or $\widehat{e}_{NL}(n_T^*)$ can be used. However, since $f_{\widehat{e}_G(n_T^*)}(x)$ is much simpler than $f_{\widehat{e}_{NL}(n_T^*)}(x)$, the former is used below. Specifically, to evaluate the right-hand side of the inequality (3.2) with $X = e$, we write:

$$
\begin{aligned}
&P\left\{\frac{|e-\widehat{e}(n_T^*)|}{e} \leqslant 2(1-\widehat{e}(n_T^*))\alpha\right\} \\
=&P\left\{\frac{e-2e\alpha}{1-2e\alpha} \leqslant \widehat{e}(n_T^*) \leqslant \frac{e+2e\alpha}{1+2e\alpha}\right\} \\
=&P\left\{\widehat{e}(n_T^*) \leqslant e+2e(1-e)\alpha+O(\alpha^2)\right\} - P\left\{\widehat{e}(n_T^*) \leqslant e-2e(1-e)\alpha+O(\alpha^2)\right\}.
\end{aligned}
\tag{B.27}
$$

To approximate (B.27), we replace $\widehat{e}(n_T^*)$ by $\widehat{e}_G(n_T^*)$. Thus, we obtain:

$$
\begin{aligned}
&P\left\{\frac{|e-\widehat{e}(n_T^*)|}{e} \leqslant 2(1-\widehat{e}(n_T^*))\alpha\right\} \\
\approx&P\left\{\widehat{e}(n_T^*) \leqslant e+2e(1-e)\alpha\right\} - P\left\{\widehat{e}(n_T^*) \leqslant e-2e(1-e)\alpha\right\} \\
=&\frac{1}{2}\left(\operatorname{erf}\left(\frac{\sqrt{n_T^*}(e+2e(1-e)\alpha-e)}{2e(1-e)}\right)\right) - \frac{1}{2}\left(\operatorname{erf}\left(\frac{\sqrt{n_T^*}(e-2e(1-e)\alpha-e)}{2e(1-e)}\right)\right) \\
=&\operatorname{erf}(\alpha\sqrt{n_T^*}) = \beta_{e,G}.
\end{aligned}
\tag{B.28}
$$

Note that in expression (3.6), the subscript G is omitted to simplify notations. □

B.4 Justification of Proposition 3.4

It follows from (3.3) and (3.6) that

$$
\beta_e = \operatorname{erf}\left(\alpha\sqrt{n_T^*(\alpha,\beta)}\right) = \operatorname{erf}\left(\frac{\alpha_e\sqrt{n_T^*(\alpha,\beta)}}{2(1-\widehat{e}(n_T^*(\alpha,\beta)))}\right). \tag{B.29}
$$

For a given (γ, δ), this expression can be rewritten as

$$\delta = \operatorname{erf}\left(\frac{\gamma\sqrt{n_e^{**}(\gamma, \delta)}}{2(1 - \widehat{e}(n_e^{**}(\gamma, \delta)))}\right), \tag{B.30}$$

where $n_e^{**}(\gamma, \delta)$ is the critical number of measurements to obtain (γ, δ)-precise estimate of e. Solving (B.30) for n_e^{**}, we obtain:

$$n_e^{**}(\gamma, \delta) = \left\lceil \left(\frac{2(1 - \widehat{e}(n_e^{**}(\gamma, \delta)))\ \operatorname{erf}^{-1}(\delta)}{\gamma}\right)^2 \right\rceil. \tag{B.31}$$

□

B.5 Additional Simulation Results of Approximation Procedure 1 for n_e^{**}

Table B.1. Parameters of randomly created machines.

Machine / Parameters	m_6	m_7	m_8	m_9	m_{10}
T_{up}	8.80	11.70	16.15	46.56	10.55
T_{down}	4.86	4.51	4.52	5.00	4.27
e	0.64	0.72	0.78	0.90	0.71
Machine / Parameters	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}
T_{up}	35.45	6.41	13.19	10.97	30.80
T_{down}	3.01	3.58	4.19	3.50	4.39
e	0.92	0.64	0.76	0.76	0.88
Machine / Parameters	m_{16}	m_{17}	m_{18}	m_{19}	m_{20}
T_{up}	8.24	8.85	17.68	21.89	6.32
T_{down}	4.56	3.31	4.97	4.58	3.97
e	0.64	0.73	0.78	0.83	0.61
Machine / Parameters	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}
T_{up}	5.98	10.27	9.71	10.88	19.79
T_{down}	3.52	3.91	3.12	3.14	4.51
e	0.63	0.72	0.76	0.78	0.81

Table B.2. Values of $\widehat{\delta}$ as a function of (γ, δ).

Machine / (γ, δ)	m_6	m_7	m_8	m_9	m_{10}
(0.1, 0.9)	0.9120	0.9169	0.9681	0.9987	0.9164
(0.1, 0.95)	0.9539	0.9625	0.9748	0.9999	0.9601
(0.05, 0.9)	0.9154	0.9248	0.9331	0.9907	0.9240
(0.05, 0.95)	0.9539	0.9648	0.9676	0.9937	0.9613
Machine / (γ, δ)	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}
(0.1, 0.9)	0.9996	0.9103	0.9529	0.9517	0.9969
(0.1, 0.95)	0.9999	0.9561	0.9591	0.9635	0.9997
(0.05, 0.9)	0.9973	0.9152	0.9295	0.9313	0.9646
(0.05, 0.95)	0.9997	0.9586	0.9665	0.9674	0.9858
Machine / (γ, δ)	m_{16}	m_{17}	m_{18}	m_{19}	m_{20}
(0.1, 0.9)	0.9141	0.9185	0.9682	0.9924	0.9128
(0.1, 0.95)	0.9561	0.9609	0.9762	0.9958	0.9572
(0.05, 0.9)	0.9173	0.9288	0.9319	0.9480	0.9199
(0.05, 0.95)	0.9595	0.9640	0.9716	0.9781	0.9593
Machine / (γ, δ)	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}
(0.1, 0.9)	0.9126	0.9204	0.9530	0.9702	0.9870
(0.1, 0.95)	0.9535	0.9572	0.9579	0.9733	0.9927
(0.05, 0.9)	0.9150	0.9230	0.9305	0.9336	0.9454
(0.05, 0.95)	0.9606	0.9631	0.9661	0.9671	0.9741

Table B.3. Values of $n_e^{**}(\gamma,\delta)$ and $\breve{n}_e^{**}(\gamma,\delta;\widehat{e}_{ini}^{\epsilon})$ as functions of (γ,δ).

Machine / $(\boldsymbol{\gamma},\boldsymbol{\delta})$	$\boldsymbol{m_6}$		$\boldsymbol{m_7}$		$\boldsymbol{m_8}$		$\boldsymbol{m_9}$		$\boldsymbol{m_{10}}$	
	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$
(0.1, 0.9)	69	85	42	56	26	40	6	30	45	60
(0.1, 0.95)	98	120	60	80	37	54	8	30	64	84
(0.05, 0.9)	275	337	168	223	104	151	21	47	180	237
(0.05, 0.95)	390	478	238	318	148	215	29	66	256	337
Machine / $(\boldsymbol{\gamma},\boldsymbol{\delta})$	$\boldsymbol{m_{11}}$		$\boldsymbol{m_{12}}$		$\boldsymbol{m_{13}}$		$\boldsymbol{m_{14}}$		$\boldsymbol{m_{15}}$	
	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$
(0.1, 0.9)	4	30	70	86	32	45	32	45	9	30
(0.1, 0.95)	5	30	99	121	45	63	45	63	12	31
(0.05, 0.9)	14	37	278	339	126	176	127	178	34	66
(0.05, 0.95)	19	51	395	484	179	252	180	253	48	93
Machine / $(\boldsymbol{\gamma},\boldsymbol{\delta})$	$\boldsymbol{m_{16}}$		$\boldsymbol{m_{17}}$		$\boldsymbol{m_{18}}$		$\boldsymbol{m_{19}}$		$\boldsymbol{m_{20}}$	
	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$
(0.1, 0.9)	69	85	41	55	27	40	17	32	81	97
(0.1, 0.95)	98	121	58	77	38	54	23	39	115	138
(0.05, 0.9)	275	338	161	216	105	153	65	105	323	387
(0.05, 0.95)	391	479	229	306	149	217	92	149	458	546
Machine / $(\boldsymbol{\gamma},\boldsymbol{\delta})$	$\boldsymbol{m_{21}}$		$\boldsymbol{m_{22}}$		$\boldsymbol{m_{23}}$		$\boldsymbol{m_{24}}$		$\boldsymbol{m_{25}}$	
	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$	n_e^{**}	$\breve{n}_e^{**}$
(0.1, 0.9)	75	90	42	56	33	46	28	41	19	34
(0.1, 0.95)	106	128	59	78	46	64	39	57	27	43
(0.05, 0.9)	298	361	165	220	129	179	109	158	75	118
(0.05, 0.95)	422	513	234	311	182	254	155	223	106	167

B.6 Justification of Proposition 3.5

To justify this proposition, we use Monte Carlo simulations to numerically evaluate β_{TP}^{ind} for various randomly generated serial lines, and check whether inequality (3.13) holds for each case. Specifically, we consider serial production lines with $M \in \{3, 5, 10, 15, 20\}$ machines. For each M, we generate 100 lines, with machine and buffer parameters randomly and equiprobably selected from the following sets:

$$T_{down,i} \in [3, 10], e_i \in [0.6, 0.95], T_{up,i} = T_{down,i}\frac{e_i}{1-e_i}, c_i \in [1, 2], i = 1, \ldots, M;$$

$$\text{Buffer capacity } N_i = \lceil r_i \max\{c_i T_{down,i}, c_{i+1} T_{down,i+1}\}\rceil, r_i \in [1, 3], i = 1, \ldots, M-1.$$

For each line, we consider $n_T^*(\alpha, \beta) \in \{108, 164, 207, 270, 384\}$, which correspond to the critical number of measurements for $\widehat{T}_{up,i}$ and $\widehat{T}_{down,i}$, $i = 1, \ldots, M$, with $(\alpha, \beta) \in \{(0.1, 0.7), (0.1, 0.8), (0.1, 0.85), (0.1, 0.90), (0.1, 0.95)\}$, respectively. Therefore, we have 2500 combinations of system parameters and numbers of up- and downtime measurements. For each of these combinations, we generate $n_T^*(\alpha, \beta)$ of up- and downtime measurements of each individual machine, and calculate the system throughput using the aggregation procedure of *Bai et al.* (2021). This process is repeated 10,000 times for each of the combinations, and β_{TP}^{ind} is evaluated as the frequency of the event in the right-hand side of (B.32), i.e.,

$$\widehat{\beta}_{TP}^{ind} = \frac{\text{number of times } \; \frac{|TP - \widehat{TP}(n_T^*(\alpha,\beta))|}{TP} \leqslant \alpha_{TP}^{ind}}{10000}. \tag{B.32}$$

As a result, we obtain that for all cases analyzed, $\widehat{\beta}_{TP}^{ind}$ exceeds $\text{erf}(\alpha\sqrt{n_T^*(\alpha, \beta)})$. In other words, Proposition 3.5 is true for all 2500 combinations considered.

□

B.7 Justification of Proposition 3.6

It follows from (3.12) and (3.13) that

$$\beta_{TP}^{ind} > \operatorname{erf}\left(\frac{\alpha_{TP}^{ind}\sqrt{n_T^*(\alpha,\beta)}}{2(1-\widehat{e}_{TP}(n_T^*(\alpha,\beta)))}\right). \tag{B.33}$$

For a given (γ, δ), this expression can be rewritten as

$$\delta > \operatorname{erf}\left(\frac{\gamma\sqrt{n_{TP}^{**}(\gamma,\delta)}}{2(1-\widehat{e}_{TP}(n_{TP}^{**}(\gamma,\delta)))}\right), \tag{B.34}$$

where $n_{TP}^{**}(\gamma,\delta)$ is the critical number of measurements to obtain (γ,δ)-precise estimate of TP. Solving (B.34) for n_{TP}^{**}, we obtain:

$$n_{TP}^{**}(\gamma,\delta) < \left\lceil\left(\frac{2(1-\widehat{e}_{TP})\operatorname{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil = \left\lceil\left(\frac{2\left(1-\frac{\widehat{TP}}{c_M}\right)\operatorname{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil. \tag{B.35}$$

□

B.8 Justification of Proposition 3.7

To simplify the notations, in this section we denote $\widehat{T}_{up}(n_T^*(\alpha,\beta))$ and $\widehat{T}_{down}(n_T^*(\alpha,\beta))$ as $\widehat{T}_{up}$ and $\widehat{T}_{down}$, respectively.

The inequalities $\frac{|\widehat{T}_{up}-Tup|}{T_{up}} \leqslant \alpha$ and $\frac{|\widehat{T}_{down}-T_{down}|}{T_{down}} \leqslant \alpha$ can be rewritten as

$$\frac{\widehat{T}_{up}}{1+\alpha} \leqslant T_{up} \leqslant \frac{\widehat{T}_{up}}{1-\alpha}, \quad \frac{\widehat{T}_{down}}{1+\alpha} \leqslant T_{down} \leqslant \frac{\widehat{T}_{down}}{1-\alpha}. \tag{B.36}$$

For T_{up} and T_{down} within the bounds (B.36), the bounds for lt, calculated using (3.15), are

$$lt_{min} \leqslant lt \leqslant lt_{max}, \tag{B.37}$$

where lt_{max} is evaluated by replacing T_{up} and T_{down} by $\frac{\widehat{T}_{up}}{1+\alpha}$ and $\frac{\widehat{T}_{down}}{1-\alpha}$; similarly, lt_{min} is evaluated by replacing T_{up} and T_{down} by $\frac{\widehat{T}_{up}}{1-\alpha}$ and $\frac{\widehat{T}_{down}}{1+\alpha}$, respectively. Thus, the bounds for the ratio $\frac{\widehat{lt}}{lt}$, where $\widehat{lt}$ is evaluated using (3.16), are given by

$$\frac{\widehat{lt}}{lt_{max}} \leqslant \frac{\widehat{lt}}{lt} \leqslant \frac{\widehat{lt}}{lt_{min}}. \tag{B.38}$$

Since both lt_{min} and lt_{max} are functions of $\alpha << 1$, using Taylor expansion, with accuracy up to $O(\alpha^2)$, we obtain:

$$\begin{aligned} \frac{\widehat{lt}}{lt_{min}} &= 1 + \left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)\left(1 + \frac{2(\widehat{\rho}+\widehat{e}-2\widehat{\rho e})}{1-\widehat{\rho}}\right)\alpha, \\ \frac{\widehat{lt}}{lt_{max}} &= 1 - \left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)\left(1 + \frac{2(\widehat{\rho}+\widehat{e}-2\widehat{\rho e})}{1-\widehat{\rho}}\right)\alpha. \end{aligned} \tag{B.39}$$

Combining both expressions in (B.39), we have:

$$-\left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)\left(1+\frac{2(\widehat{\rho}+\widehat{e}-2\widehat{\rho e})}{1-\widehat{\rho}}\right)\alpha \leqslant \frac{\widehat{lt}-lt}{lt} \leqslant \left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)\left(1+\frac{2(\widehat{\rho}+\widehat{e}-2\widehat{\rho e})}{1-\widehat{\rho}}\right)\alpha$$
$$\iff \frac{|\widehat{lt}-lt|}{lt} \leqslant \left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)\left(1+\frac{2(\widehat{\rho}+\widehat{e}-2\widehat{\rho e})}{1-\widehat{\rho}}\right)\alpha. \tag{B.40}$$

Thus, α_{lt}^{ind} with accuracy up to $O(\alpha^2)$ is given by:

$$\alpha_{lt}^{ind} = \left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)\left(1+\frac{2(\widehat{\rho}+\widehat{e}-2\widehat{\rho e})}{1-\widehat{\rho}}\right)\alpha. \tag{B.41}$$

□

B.9 Justification of Proposition 3.8

The condition (3.19) can be rewritten as

$$(1-\widehat{e})\widehat{e} > \frac{\tau}{4\widehat{T}_{down}}. \tag{B.42}$$

We prove by contradiction that this inequality implies $\frac{\alpha_{lt}^{ind}}{\alpha} > 1$. Assume (B.42) holds and $\frac{\alpha_{lt}^{ind}}{\alpha} \leqslant 1$. Then,

$$\frac{\alpha_{lt}^{ind}}{\alpha} \leqslant 1 \iff \frac{(1-\widehat{e})(\widehat{\rho}+\widehat{e}-2\widehat{e}\widehat{\rho})}{(1-\widehat{\rho})^2} \leqslant \frac{\tau}{4\widehat{T}_{down}}. \tag{B.43}$$

Denote the expression $\frac{(1-\widehat{e})(\widehat{\rho}+\widehat{e}-2\widehat{e}\widehat{\rho})}{(1-\widehat{\rho})^2}$ in (B.43) as $K(\widehat{e},\widehat{\rho})$, and observe that

$$\frac{\partial K}{\partial \widehat{\rho}} = \frac{(1-\widehat{e})(1+\widehat{\rho}-2\widehat{e}\widehat{\rho})}{(1-\widehat{\rho})} > 0. \tag{B.44}$$

Inequality (B.44) indicates that K is an increasing function of $\widehat{\rho}$. In addition, for $\widehat{\rho} = 0$, $K = (1-\widehat{e})\widehat{e}$. Thus, for $\widehat{\rho} \geqslant 0$, we have

$$K(\widehat{e}, \widehat{\rho}) \geqslant (1-\widehat{e})\widehat{e}. \tag{B.45}$$

Combining the above with the assumption (B.42), we obtain:

$$K(\widehat{e}, \widehat{\rho}) \geqslant (1-\widehat{e})\widehat{e} > \frac{\tau}{4\widehat{T}_{down}}, \tag{B.46}$$

which contradicts with (B.43).

On the other hand, if $\frac{\alpha_{lt}^{ind}}{\alpha} > 1$, i.e., $K(\widehat{e}, \widehat{\rho}) = \frac{(1-\widehat{e})(\widehat{\rho}+\widehat{e}-2\widehat{e}\widehat{\rho})}{(1-\widehat{\rho})^2} > \frac{\tau}{4\widehat{T}_{down}}$, then $(1-\widehat{e})\widehat{e} > \frac{\tau}{4\widehat{T}_{down}}$ as well.

□

B.10 Justification of Proposition 3.9

To simplify notations, in this section we denote $n_T^*(\alpha, \beta)$ as n. Similar to the approach adopted in Appendix B, we use Taylor expansion to obtain a linear approximation of $\widehat{lt}(n)$ with respect to its arguments $\widehat{T}_{up}(n)$ and $\widehat{T}_{down}(n)$, denoted as $\widehat{lt}_L(n)$:

$$\begin{aligned}\widehat{lt}_L(n) = lt &+ \left.\frac{\partial \widehat{lt}(n)}{\partial \widehat{T}_{up}(n)}\right|_{\widehat{T}_{up}(n)=T_{up}, \widehat{T}_{down}(n)=T_{down}} \left(\widehat{T}_{up}(n) - T_{up}\right) \\ &+ \left.\frac{\partial \widehat{lt}(n)}{\partial \widehat{T}_{down}(n)}\right|_{\widehat{T}_{up}(n)=T_{up}, \widehat{T}_{down}(n)=T_{down}} \left(\widehat{T}_{down}(n) - T_{down}\right).\end{aligned} \tag{B.47}$$

Since $\widehat{T}_{up}(n)$ and $\widehat{T}_{down}(n)$ can be approximated by Gaussian distributions $\mathcal{N}(T_{up}, \frac{T_{up}^2}{n})$

and $\mathcal{N}(T_{down}, \frac{T^2_{down}}{n})$, respectively, $\widehat{lt}_L(n)$ also can be approximated by a Gaussian random variable, $\widehat{lt}_G(n)$, with expected value equal to lt, and variance

$$\begin{aligned} var\left(\widehat{lt}_G(n)\right) &= \left(\frac{\partial \widehat{lt}(n)}{\partial \widehat{T}_{up}(n)}\Bigg|_{\widehat{T}_{up}(n)=T_{up}, \widehat{T}_{down}(n)=T_{down}}\right)^2 var(\widehat{T}_{up}(n)) \\ &\quad \left(\frac{\partial \widehat{lt}(n)}{\partial \widehat{T}_{down}(n)}\Bigg|_{\widehat{T}_{up}(n)=T_{up}, \widehat{T}_{down}(n)=T_{down}}\right)^2 var(\widehat{T}_{down}(n)) \\ &= \frac{1}{n}(lt-1)^2 \frac{(\rho + e - 2e\rho)^2 + (1 + e - 2e\rho)^2}{(1-\rho)^2}. \end{aligned} \tag{B.48}$$

Therefore, the Gaussian approximation of β^{ind}_{lt} is given by

$$\begin{aligned} \beta^{ind}_{lt} &= P\left(\frac{|\widehat{lt}(n) - lt|}{lt} \leqslant \alpha^{ind}_{lt}\right) \approx P\left(\frac{|\widehat{lt}_G(n) - lt|}{lt} \leqslant \alpha^{ind}_{lt}\right) \\ &= P\left(\widehat{lt}_G(n) \leqslant (1+\alpha^{ind}_{lt})lt\right) - P\left(\widehat{lt}_G(n) \leqslant (1-\alpha^{ind}_{lt})lt\right) \\ &= \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{(1+\alpha^{ind}_{lt})lt - lt}{\sqrt{2var\left(\widehat{lt}_G(n)\right)}}\right)\right) - \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{(1-\alpha^{ind}_{lt})lt - lt}{\sqrt{2var\left(\widehat{lt}_G(n)\right)}}\right)\right) \\ &\approx \mathrm{erf}\left(\alpha A \sqrt{\frac{n}{2}}\right), \end{aligned} \tag{B.49}$$

where $A = \sqrt{\frac{(1+\widehat{\rho}+2\widehat{e}-4\widehat{e}\widehat{\rho})}{(\widehat{\rho}+\widehat{e}-2\widehat{e}\widehat{\rho})^2+(1+\widehat{e}-2\widehat{e}\widehat{\rho})^2}}$.

□

B.11 Justification of Proposition 3.10

As mentioned in Appendix B, $\widehat{lt}(n^*_T(\alpha, \beta))$ can be approximated by a Gaussian random variable $\widehat{lt}_G(n_T^*(\alpha,\beta)) \sim \mathcal{N}\Big(lt, \frac{1}{n_T^*(\alpha,\beta)}(lt-1)^2\frac{(\rho+e-2e\rho)^2+(1+e-2e\rho)^2}{(1-\rho)^2}\Big)$. Thus, for any given accuracy γ, the Gaussian approximation of its probability δ is given by

$$\begin{aligned}\delta =& P\left(\frac{|\widehat{lt}(n_{LT}^{**}(\gamma,\delta))-lt|}{lt}\leqslant\gamma\right)\approx P\left(\frac{|\widehat{lt}_G(n_{LT}^{**}(\gamma,\delta))-lt|}{lt}\leqslant\gamma\right)\\ =&\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{(1+\gamma)lt-lt}{\sqrt{2var\left(\widehat{lt}_G(n_{LT}^{**}(\gamma,\delta))\right)}}\right)\right)\\ &-\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{(1-\gamma)lt-lt}{\sqrt{2var\left(\widehat{lt}_G(n_{LT}^{**}(\gamma,\delta))\right)}}\right)\right)\\ \approx&\operatorname{erf}\left(\sqrt{\frac{n_{LT}^{**}(\gamma,\delta)(1-\widehat{\rho})^2}{2\left((\widehat{\rho}+\widehat{e}-2\widehat{e}\widehat{\rho})^2+(1+\widehat{e}-2\widehat{e}\rho)^2\right)}}\left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)\delta\right).\end{aligned} \tag{B.50}$$

Solving for $n_{LT}^{**}(\gamma,\delta)$, we obtain:

$$n_{LT}^{**}(\gamma,\delta)=\left\lceil 2R\left(\frac{\widehat{lt}-1}{\widehat{lt}}\right)^2\left(\frac{\operatorname{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil, \tag{B.51}$$

where

$$R=\frac{(\widehat{e}+\widehat{\rho}-2\widehat{\rho}\widehat{e})^2+(1+\widehat{e}-2\widehat{\rho}\widehat{e})^2}{(1-\widehat{\rho})^2}. \tag{B.52}$$

□

B.12 Justification of Proposition 3.11

Using the same arguments as in Appendix B and substituting WIP instead of

lt, we obtain:

$$\frac{\widehat{WIP}}{WIP_{max}} \leqslant \frac{\widehat{WIP}}{WIP} \leqslant \frac{\widehat{WIP}}{WIP_{min}}. \tag{B.53}$$

Similar to Appendix B, taking Taylor expansion, with accuracy up to $O(\alpha^2)$ we have:

$$\begin{aligned} \frac{\widehat{WIP}}{WIP_{min}} &= 1 + \frac{1 + 2\widehat{e} + \widehat{\rho} - 4\widehat{e}\widehat{\rho}}{1 - \widehat{\rho}}\alpha, \\ \frac{\widehat{WIP}}{WIP_{max}} &= 1 - \frac{1 + 2\widehat{e} + \widehat{\rho} - 4\widehat{e}\widehat{\rho}}{1 - \widehat{\rho}}\alpha. \end{aligned} \tag{B.54}$$

Combining both expressions in (B.54), we obtain:

$$-\frac{1 + 2\widehat{e} + \widehat{\rho} - 4\widehat{e}\widehat{\rho}}{1 - \widehat{\rho}}\alpha \leqslant \frac{\widehat{WIP} - WIP}{WIP} \leqslant \frac{1 + 2\widehat{e} + \widehat{\rho} - 4\widehat{e}\widehat{\rho}}{1 - \widehat{\rho}}\alpha \tag{B.55}$$

$$\iff \frac{|\widehat{WIP} - WIP|}{WIP} \leqslant \frac{1 + 2\widehat{e} + \widehat{\rho} - 4\widehat{e}\widehat{\rho}}{1 - \widehat{\rho}}\alpha. \tag{B.56}$$

Thus, with accuracy up to $O(\alpha^2)$, α_{WIP}^{ind} is given by

$$\alpha_{WIP}^{ind} = \frac{1 + 2\widehat{e} + \widehat{\rho} - 4\widehat{e}\widehat{\rho}}{1 - \widehat{\rho}}\alpha. \tag{B.57}$$

□

B.13 Justification of Proposition 3.12

Follows exactly the Justification of Proposition 3.9. The only difference is that in the expressions (B.47)-(B.49), lt, $\widehat{lt}$, $(lt-1)^2$ should be substituted by WIP, $\widehat{WIP}$, $(WIP)^2$.

□

B.14 Justification of Proposition 3.13

Follows Appendix B with lt, $\widehat{lt}$, $(lt-1)^2$ substituted by WIP, $\widehat{WIP}$, $(WIP)^2$ and taking into the account that

$$var(\widehat{WIP}_G(n^{**}_{WIP}(\alpha,\beta))) = \frac{1}{n^*(\alpha,\beta)}(WIP)^2\frac{(\rho+e-2e\rho)^2+(1+e-2e\rho)^2}{(1-\rho)^2}. \tag{B.58}$$

As a result, we obtain:

$$n^{**}_{WIP}(\gamma,\delta) = \left\lceil 2R\left(\frac{\mathrm{erf}^{-1}(\delta)}{\gamma}\right)^2\right\rceil, \tag{B.59}$$

where R is given by (B.52).

□

APPENDIX C

Appendix of Chapter 4

C.1 Justification of Proposition 4.1

The Gaussian approximation of $\widehat{q}(n)$ is $\mathcal{N}\left(q, \frac{(1-q)q}{n}\right)$ (see *Feller* (2015)). Therefore,

$$\begin{aligned}
\beta_q &= P\left(\frac{|q-\widehat{q}(n)|}{q} \leqslant \alpha_q\right) = P\left((1-\alpha_q)q \leqslant \widehat{q}(n) \leqslant (1+\alpha_q)q\right) \\
&\approx \frac{1}{2}\left(1+\operatorname{erf}\left(\frac{(1+\alpha_q)q}{\sqrt{2}\sqrt{\frac{(1-q)q}{n}}}\right)\right) - \frac{1}{2}\left(1+\operatorname{erf}\left(\frac{(1-\alpha_q)q}{\sqrt{2}\sqrt{\frac{(1-q)q}{n}}}\right)\right) \\
&= \operatorname{erf}\left(\frac{\alpha_q\sqrt{n}}{\sqrt{2}\sqrt{\frac{1-q}{q}}}\right) \approx \operatorname{erf}\left(\frac{\alpha_q\sqrt{n_q^*(\alpha_q,\beta_q)}}{\sqrt{2}\sqrt{\frac{1-\widehat{q}}{\widehat{q}}}}\right).
\end{aligned} \tag{C.1}$$

Solving for $n_q^*(\alpha_q, \beta_q)$, we obtain a Gaussian approximation of the critical number as follows:

$$n_q^*(\alpha_q, \beta_q) = \left\lceil 2\left(\frac{1-\widehat{q}}{\widehat{q}}\right)\left(\frac{\mathrm{erf}^{-1}(\beta_q)}{\alpha_q}\right)^2\right\rceil. \tag{C.2}$$

□

C.2 Justification of Proposition 4.2

For $\widehat{q}_i(n_q^*(\alpha_q, \beta_q))$, $i = 1, \ldots, k$, satisfying $(1-\alpha_q)q_i \leqslant \widehat{q}_i(n_q^*(\alpha_q, \beta_q)) \leqslant (1+\alpha_q)q_i$, we have:

$$\begin{aligned} &(1-\alpha_q)^k \prod_{i=1}^{k} q_i \leqslant \prod_{i=1}^{k} \widehat{q}_i \leqslant (1+\alpha_q)^k \prod_{i=1}^{k} q_i \\ \iff &(1-k\alpha_q + O(\alpha_q^2)) \prod_{i=1}^{k} q_i \leqslant \prod_{i=1}^{k} \widehat{q}_i \leqslant (1+k\alpha_q + O(\alpha_q^2)) \prod_{i=1}^{k} q_i. \end{aligned} \tag{C.3}$$

Thus,

$$\begin{aligned} &(1-k\alpha_q + O(\alpha_q^2))TP \prod_{i=1}^{k} q_i \leqslant TP \prod_{i=1}^{k} \widehat{q}_i \leqslant (1+k\alpha_q + O(\alpha_q^2))TP \prod_{i=1}^{k} q_i \\ \iff &(1-k\alpha_q + O(\alpha_q^2))TP_q \leqslant \widehat{TP}_q \leqslant (1+k\alpha_q + O(\alpha_q^2))TP_q \\ \iff &\frac{|\widehat{TP}_q - TP_q|}{TP_q} \leqslant k\alpha_q + O(\alpha_q^2). \end{aligned} \tag{C.4}$$

Therefore, with accuracy up to $O(\alpha_q^2)$,

$$\alpha_{TP_q}^{ind} = k\alpha_q. \tag{C.5}$$

□

C.3 Justification of Proposition 4.3

To simplify the notations, in this subsection we denote $n_q^*(\alpha_q, \beta_q)$ as n, and $\prod_{i=1}^{k} \widehat{q}_i(n) = \widehat{Q}(n)$.

Since the Gaussian approximation of $\widehat{q}_i(n)$, $i = 1, \ldots, k$, is $\mathcal{N}\left(q_i, \frac{(1-q_i)q_i}{n}\right)$, the expected value and variance of $\widehat{Q}(n)$ are:

$$E(\widehat{Q}(n)) = \prod_{i=1}^{k} q_i =: Q, \tag{C.6}$$

$$\begin{aligned} var(\widehat{Q}(n)) =& E(\widehat{Q}(n)^2) - E(\widehat{Q}(n))^2 \\ =& \prod_{i=1}^{k} \left(var(\widehat{q}_i(n)) + E(\widehat{q}_i(n))^2\right) - \prod_{i=1}^{k} E(\widehat{q}_i(n))^2 \\ =& \prod_{i=1}^{k} \left(\frac{(1-q_i)q_i}{n} + q_i^2\right) - \prod_{i=1}^{k} q_i^2. \end{aligned} \tag{C.7}$$

If $k \geqslant 2$, with accuracy up to $O(\frac{1}{n^k})$, this expression becomes:

$$var(\widehat{Q}(n)) = \frac{1}{n} \sum_{i=1}^{k} \frac{\prod_{j=1}^{k} q_j^2}{q_i^2} q_i(1-q_i) = \frac{1}{n} \sum_{i=1}^{k} Q^2 \left(\frac{1}{q_i} - 1\right). \tag{C.8}$$

The last term in (C.8) has the following upper bound (see the proof below):

$$\sum_{i=1}^{k} Q^2 \left(\frac{1}{q_i} - 1\right) < \frac{1}{4}. \tag{C.9}$$

This leads to the following:

$$\beta^{ind}_{TP_q} = P\left(\frac{|\widehat{TP}_q(n) - TP_q|}{TP_q} \leqslant \alpha^{ind}_{TP_q}\right) = P\left(\frac{|TP\prod_{i=1}^{k}\widehat{q}_i - TP\prod_{i=1}^{k}q_i|}{TP\prod_{i=1}^{k}q_i} \leqslant \alpha^{ind}_{TP_q}\right)$$
$$= P\left(\frac{|\widehat{Q}(n) - Q|}{Q} \leqslant \alpha^{ind}_{TP_q}\right). \tag{C.10}$$

Keeping in mind that $\widehat{Q}(n)$ can be approximated by a Gaussian distribution $\mathcal{N}\left(Q, \frac{1}{n}\sum_{i=1}^{k}Q^2\left(\frac{1}{q_i}-1\right)\right)$ and using (C.9) in (C.10), we obtain:

$$\beta^{ind}_{TP_q} = P\left(\frac{|\widehat{Q}(n) - Q|}{Q} \leqslant \alpha^{ind}_{TP_q}\right)$$
$$\approx \frac{1}{2}\left(1 + \text{erf}\left(\frac{(1+\alpha^{ind}_{TP_q})Q - Q}{\sqrt{2var(\widehat{Q}(n))}}\right)\right) - \frac{1}{2}\left(1 + \text{erf}\left(\frac{(1-\alpha^{ind}_{TP_q})Q - Q}{\sqrt{2var(\widehat{Q}(n))}}\right)\right)$$
$$> \text{erf}\left(\frac{\alpha^{ind}_{TP_q}Q}{\sqrt{2}\sqrt{\frac{1}{4n}}}\right) = \text{erf}\left(\sqrt{2n}k\alpha_q Q\right). \tag{C.11}$$

For $Q > 0.5$, we have:

$$\beta^{ind}_{TP_q} > \text{erf}\left(\frac{k\alpha_q}{\sqrt{2}}\sqrt{n^*_q(\alpha_q, \beta_q)}\right), \tag{C.12}$$

which coincides with the statement of Proposition 4.3.

To prove (C.9), introduce the notation:

$$S_k(q_1, \ldots, q_k) = \sum_{i=1}^{k} Q^2 \left(\frac{1}{q_i} - 1 \right). \tag{C.13}$$

The partial derivative $\frac{\partial S_k}{\partial q_i}$ is

$$\frac{\partial S_k}{\partial q_i} = \sum_{l \neq i} 2q_i \left(\prod_{h \neq l, h \neq i} g_h^2 \right) g_l(1 - g_l) + \left(\prod_{l \neq i} q_l^2 \right) q_i(1 - 2q_i), \text{ for } i = 1, \ldots, k. \tag{C.14}$$

Setting $\frac{\partial S_k}{\partial q_i} = 0$, we obtain:

$$\sum_{l \neq i} 2q_i^* \left(\prod_{h \neq l, h \neq i} q_h^* \right) + \prod_{h \neq i} q_h^* = 2k \left(\prod_{i=1}^{k} q_i^* \right), \tag{C.15}$$

where $q_i^* = \arg(\max S_k)$, $i = 1, \ldots, k$. Since the right-hand side of (C.15) remains the same for $\forall j \neq i$, we conclude that $q_i^* = q_j^* =: q^*$. Therefore, (C.15) can be rewritten as

$$2(k-1)(q^*)^{k-1} + (q^*)^{k-1} = 2k(q^*)^k, \tag{C.16}$$

implying that

$$q^* = \frac{2k-1}{2k}, \tag{C.17}$$

and, therefore,

$$S_k^* = \max S_k = \frac{1}{2} \left(\frac{2k-1}{2k} \right)^{2k-1}. \tag{C.18}$$

Since S_k^* is a decreasing function of k, we obtain:

$$\max_{k \geqslant 2} S_k^* = S_2^* = \frac{27}{128} < \frac{1}{4}, \tag{C.19}$$

which proves (C.9).

□

C.4 Justification of Proposition 4.4

From Appendix C, $\widehat{Q}(n_q^*(\alpha_q, \beta_q)) = \prod_{i=1}^k \widehat{q}_i(n_q^*(\alpha_q, \beta_q))$ can be approximated by a Gaussian distribution $\mathcal{N}\left(Q, \frac{1}{n_q^*(\alpha_q,\beta_q)} \sum_{i=1}^k Q^2 \left(\frac{1}{q_i} - 1\right)\right)$, where $Q = \prod_{i=1}^k q_i$. Thus, for any given accuracy γ, the Gaussian approximation of its corresponding probability δ is given by

$$\begin{aligned}
\delta &= P\left(\frac{|\widehat{TP}_q(n_{TP_q}^{**}(\gamma,\delta)) - TP_q|}{TP_q} \leqslant \gamma\right) = P\left(\frac{|\widehat{Q}(n_{TP_q}^{**}(\gamma,\delta)) - Q|}{Q} \leqslant \gamma\right) \\
&\approx \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{(1+\gamma)Q - Q}{\sqrt{2var\left(\widehat{Q}(n_{TP_q}^{**}(\gamma,\delta))\right)}}\right)\right) \\
&\quad - \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{(1-\gamma)Q - Q}{\sqrt{2var\left(\widehat{Q}(n_{TP_q}^{**}(\gamma,\delta))\right)}}\right)\right) \\
&= \operatorname{erf}\left(\frac{\gamma\sqrt{n_{TP_q}^{**}(\gamma,\delta)}}{\sqrt{2}\sqrt{\sum_{i=1}^k \frac{1-q_i}{q_i}}}\right) \approx \operatorname{erf}\left(\frac{\gamma\sqrt{n_{TP_q}^{**}(\gamma,\delta)}}{\sqrt{2}\sqrt{\sum_{i=1}^k \frac{1-\widehat{q}_i}{\widehat{q}_i}}}\right).
\end{aligned} \tag{C.20}$$

Solving for $n_{TP_q}^{**}(\gamma,\delta)$, we obtain the Gaussian approximation of the critical number as follows:

$$n_{TP_q}^{**}(\gamma,\delta) = \left\lceil 2\left(\sum_{i=1}^k \frac{1-\widehat{q}_i}{\widehat{q}_i}\right)\left(\frac{\operatorname{erf}^{-1}(\delta)}{\gamma}\right)^2 \right\rceil. \tag{C.21}$$

□

APPENDIX D

Appendix of Chapter 5

D.1 Justification of Proposition 5.1

Similar to A, the estimate $\widehat{T}_{up}(n)$ can be approximated by a Gaussian random variable $\widehat{T}_{up,G}(n) \sim \mathcal{N}(T_{up}, \frac{T_{up}^2}{n})$. Thus, according to *Tsagris et al.* (2014), the mean of the random variable $|\widehat{T}_{up,G}(n) - T_{up}|$ is $\sqrt{\frac{2T_{up}^2}{\pi n}}$. By Markov's inequality, we have

$$
\begin{aligned}
P\left(|\widehat{T}_{up,G}(n) - T_{up}| \geqslant a\right) &\leqslant \frac{E(|\widehat{T}_{up,G}(n) - T_{up}|)}{a} \\
\Longrightarrow P\left(|\widehat{T}_{up,G}(n) - T_{up}| \leqslant a\right) &\geqslant 1 - \frac{E(|\widehat{T}_{up,G}(n) - T_{up}|)}{a} \\
\Longrightarrow P\left(|\widehat{T}_{up,G}(n) - T_{up}| \leqslant a\right) &\geqslant 1 - \frac{1}{a}\sqrt{\frac{2T_{up}^2}{\pi n}},
\end{aligned}
$$

Take $a = \alpha T_{up}$, we have $\beta = 1 - \frac{1}{\alpha T_{up}}\sqrt{\frac{2T_{up}^2}{\pi n^*}} = 1 - \frac{1}{\alpha}\sqrt{\frac{2}{\pi n^*}}$, where $n^* = n^*_{T,M}(\alpha, \beta)$. Solving for $n^*_{T,M}(\alpha, \beta)$, we obtain

$$n^*_{T,M}(\alpha,\beta) = \left\lceil \frac{2}{\pi\alpha^2(1-\beta)^2} \right\rceil.$$

□

D.2 Justification of Proposition 5.2

Similar to the justification of 5.1. □

D.3 Proof of Theorem 5.3

The mean and variance of the estimate $\widehat{T}_{up}(n)$ are T_{up} and $\frac{T^2_{up}}{n}$, respectively. Therefore, by Chebyshev's inequality, we have

$$P\left(|\widehat{T}_{up}(n^*) - T_{up}| \geqslant k\frac{T_{up}}{\sqrt{n^*}}\right) \leqslant \frac{1}{k^2}$$
$$\Longrightarrow P\left(\frac{|\widehat{T}_{up}(n^*) - T_{up}|}{T_{up}} \leqslant \frac{k}{\sqrt{n^*}}\right) \geqslant 1 - \frac{1}{k^2},$$

where $n^* = n^*_{T,C}(\alpha,\beta)$. Take $\alpha = \frac{k}{\sqrt{n^*_{T,C}(\alpha,\beta)}}$, we have $\beta = 1 - \frac{1}{k^2}$. Solving for $n^*_{T,C}(\alpha,\beta)$, we obtain

$$n^*_{T,C}(\alpha,\beta) = \left\lceil \frac{1}{\alpha^2(1-\beta)} \right\rceil.$$

□

D.4 Proof of Theorem 5.4

Similar to the justification of 5.3. □

APPENDIX E

Appendix of Chapter 6

E.1 Proof of Theorem 6.1

Consider an unreliable machine with cycle overrun, defined by $\{\tau, T_{up}, T_{down}, p_{OR}, k_{OR}\}$, at the end of its N-th up- and downtime realization. Denote the number of parts produced during this time period as K. This time period equals to $\sum_{i=1}^{N}(t_{up,i} + t_{down,i})$, and the machine's total uptime is $\sum_{i=1}^{N} t_{up,i}$. Denote the duration of the processing time of k-th part as $\tau_{total,k} = \tau + \tau_{OR,k}$, where $\tau_{OR,k}$ is the realization of the k-th overrun, the mean value of which is T_{OR}. Since there are K parts produced, $\sum_{i=1}^{N} t_{up,i} \geqslant \sum_{k=1}^{K} \tau_{total,k}$. Let $\sum_{i=1}^{N} t_{up,i} = \sum_{k=1}^{K} \tau_{total,k} + \check{\tau}$, where $\check{\tau} < \tau_{total,K+1}$.

Assume that $N \to \infty$, then, obviously, $K \to \infty$ as well. Hence, the stand-alone throughput of the machine with cycle overrun can be evaluated as follows:

$$\begin{aligned} SAT &= \lim_{N\to\infty, K\to\infty} \frac{K}{\sum_{i=1}^{N}(t_{up,i}+t_{down,i})} \\ &= \lim_{N\to\infty, K\to\infty} \left[\left(\frac{K}{\sum_{k=1}^{K}\tau_{total,k}+\check{\tau}} \right) \left(\frac{\sum_{i=1}^{N} t_{up,i}}{\sum_{i=1}^{N}(t_{up,i}+t_{down,i})} \right) \right] \\ &= \left(\lim_{K\to\infty} \frac{K}{\sum_{k=1}^{K}\tau_{total,k}+\check{\tau}} \right) \left(\lim_{N\to\infty} \frac{\sum_{i=1}^{N} t_{up,i}}{\sum_{i=1}^{N}(t_{up,i}+t_{down,i})} \right). \end{aligned} \tag{E.1}$$

The first term in the last row of (E.1) can be bounded as follows:

$$\lim_{K\to\infty} \frac{K}{\sum_{k=1}^{K+1}\tau_{total,k}} \leqslant \lim_{K\to\infty} \frac{K}{\sum_{k=1}^{K}\tau_{total,k}+\check{\tau}} \leqslant \lim_{K\to\infty} \frac{K}{\sum_{k=1}^{K}\tau_{total,k}}. \tag{E.2}$$

The first term in (E.2) can be represented as

$$\begin{aligned} \lim_{K\to\infty} \frac{K}{\sum_{k=1}^{K+1}\tau_{total,k}} &= \lim_{K\to\infty} \frac{K}{K+1} \frac{K+1}{\sum_{k=1}^{K+1}\tau_{total,k}} \\ &= \left(\lim_{K\to\infty} \frac{K}{K+1} \right) \left(\lim_{K\to\infty} \frac{K+1}{\sum_{k=1}^{K+1}\tau_{total,k}} \right) \\ &= \lim_{K\to\infty} \frac{K+1}{\sum_{k=1}^{K+1}\tau_{total,k}}. \end{aligned} \tag{E.3}$$

Thus, using the last expression in (E.3) and the strong law of large numbers applied to the first and third terms of (E.2), we obtained:

$$\lim_{K\to\infty} \frac{K+1}{\sum_{k=1}^{K+1}\tau_{total,k}} = \lim_{K\to\infty} \frac{K}{\sum_{k=1}^{K}\tau_{total,k}} = \frac{1}{\tau+T_{OR}}. \tag{E.4}$$

Therefore, the middle term of (E.2) is also given by

$$\lim_{K\to\infty} \frac{K}{\sum_{k=1}^{K} \tau_{total,k} + \check{\tau}} = \frac{1}{\tau + T_{OR}}. \tag{E.5}$$

As far as the second term in (E.1) is concerned, by the strong law of large numbers we have:

$$\lim_{N\to\infty} \frac{\sum_{i=1}^{N} t_{up,i}}{\sum_{i=1}^{N}(t_{up,i} + t_{down,i})} = \frac{\lim\limits_{N\to\infty} \frac{\sum_{i=1}^{N} t_{up,i}}{N}}{\lim\limits_{N\to\infty} \frac{\sum_{i=1}^{N} t_{up,i}}{N} + \lim\limits_{N\to\infty} \frac{\sum_{i=1}^{N} t_{down,i}}{N}} = \frac{T_{up}}{T_{up} + T_{down}}. \tag{E.6}$$

Thus, combining (E.5) and (E.6), and assuming T_{up} and T_{down} are in seconds, we obtain:

$$\begin{aligned} SAT =& \frac{1}{\tau + T_{OR}} \frac{T_{up}}{T_{up} + T_{down}} \text{ parts/second} \\ =& \frac{3600}{\tau + T_{OR}} \frac{T_{up}}{T_{up} + T_{down}} \text{ parts/hour.} \end{aligned} \tag{E.7}$$

□

E.2 Proof of Theorem 6.2

In the following, we prove that $\frac{T_{down}}{T_{up}} < p_{OR}k_{OR}$ is a necessary and sufficient condition for the overrun-reduced machine having a larger SAT than that of the downtime-reduced one.

To show necessity, assume that the overrun-reduced machine has a larger SAT

than that of the downtime-reduced one. In this case, we have:

$$\begin{aligned}
&\frac{3600}{\tau + rp_{OR}k_{OR}\tau}\frac{T_{up}}{T_{up}+T_{down}} > \frac{3600}{\tau + p_{OR}k_{OR}\tau}\frac{T_{up}}{T_{up}+rT_{down}} \\
&\Leftrightarrow (\tau + rp_{OR}k_{OR}\tau)(T_{up}+T_{down}) < (\tau + p_{OR}k_{OR}\tau)(T_{up}+rT_{down}) \\
&\Leftrightarrow (1-r)\tau T_{down} < (1-r)p_{OR}k_{OR}\tau T_{up} \\
&\Leftrightarrow \frac{T_{down}}{T_{up}} < p_{OR}k_{OR}.
\end{aligned} \tag{E.8}$$

Note that the chain of inequalities in (E.8) is bi-directional, thus, $\frac{T_{down}}{T_{up}} < p_{OR}k_{OR}$ is both necessary and sufficient.

Similarly, it can be shown that the downtime-reduced machine has a larger SAT than that of the overrun-reduced one, i.e., $\frac{3600}{\tau+rp_{OR}k_{OR}\tau}\frac{T_{up}}{T_{up}+T_{down}} < \frac{3600}{\tau+p_{OR}k_{OR}\tau}\frac{T_{up}}{T_{up}+rT_{down}}$ if and only if $\frac{T_{down}}{T_{up}} > p_{OR}k_{OR}$ holds. □

E.3 Proof of Theorem 6.3

Consider the equations:

$$\begin{aligned}
T_{OR}^c &= \alpha\Gamma(1+\frac{1}{\beta}), \\
CV &= \sqrt{\frac{\Gamma(1+\frac{2}{\beta})}{\Gamma^2(1+\frac{1}{\beta})} - 1}\,.
\end{aligned} \tag{E.9}$$

Introduce function

$$G(\beta) = \frac{\Gamma(1+\frac{2}{\beta})}{\Gamma^2(1+\frac{1}{\beta})} \tag{E.10}$$

and consider its derivative with respect to β:

$$G'(\beta) = \frac{2\Gamma(1+\frac{2}{\beta})\Gamma'(1+\frac{1}{\beta}) - 2\Gamma(1+\frac{1}{\beta})\Gamma'(1+\frac{2}{\beta})}{\beta^2\Gamma^3(1+\frac{1}{\beta})}. \tag{E.11}$$

As one can see, the denominator of (E.11) is positive. To prove that the numerator is negative (i.e., that $G(\beta)$ is monotonically decreasing) consider what is referred to as the digamma function, $\psi(x)$, and its derivative, $\psi'(x)$, given by (see *Abramowitz and Stegun* (1948)):

$$\begin{aligned} \psi(x) &= \frac{\Gamma'(x)}{\Gamma(x)}, \\ \psi'(x) &= -\int_0^1 \frac{t^{x-1}}{1-t}\ln t dt > 0 \text{ for } x > 0. \end{aligned} \tag{E.12}$$

Thus, for $x > 0$, $\psi(x)$ is monotonically increasing, i.e., $\psi(1+\frac{2}{\beta}) > \psi(1+\frac{1}{\beta})$. In other words,

$$\begin{aligned} &\frac{\Gamma'(1+\frac{1}{\beta})}{\Gamma(1+\frac{1}{\beta})} < \frac{\Gamma'(1+\frac{2}{\beta})}{\Gamma(1+\frac{2}{\beta})} \\ \iff &\Gamma(1+\frac{2}{\beta})\Gamma'(1+\frac{1}{\beta}) < \Gamma(1+\frac{1}{\beta})\Gamma'(1+\frac{2}{\beta}) \\ \iff &\Gamma(1+\frac{2}{\beta})\Gamma'(1+\frac{1}{\beta}) - \Gamma(1+\frac{1}{\beta})\Gamma'(1+\frac{2}{\beta}) < 0. \end{aligned} \tag{E.13}$$

Therefore, $G(\beta)$ and, hence, $CV(\beta)$ are decreasing functions of $\beta > 0$. In particular, $CV = 1$ for $\beta = 1$ and tends to 0 when $\beta \to \infty$. Due to monotonicity of $CV(\beta)$, we conclude that for $CV < 1$, β is unique in the range of $\beta \in (1, +\infty)$.

Since β is unique, using (E.9), we conclude that α is unique as well in the range of $\alpha \in (0, +\infty)$. □

E.4 Justification of Proposition 6.5

Denote the estimate of mean conditional cycle overrun based on $n^*_{OR}(\alpha, \beta)$ cycle overrun observations as $\widehat{T}^c_{OR}(n^*_{OR})$, i.e.,

$$\widehat{T}^c_{OR}(n^*_{OR}) = \frac{\sum_{i=1}^{n^*_{OR}} t_{OR,i}}{n^*_{OR}}, \tag{E.14}$$

where $t_{OR,i}$ is the i-th cycle overrun duration. Thus, by (6.19), we have $\widehat{T}^c_{OR}(n^*_{OR}) = \widehat{k}_{OR}(n^*_{OR})\tau$, and

$$\begin{aligned} P\left\{\frac{|\widehat{k}_{OR}(n^*_{OR}) - k|}{k} \leqslant \alpha\right\} &= P\left\{\frac{|\widehat{k}_{OR}(n^*_{OR})\tau - k\tau|}{k\tau} \leqslant \alpha\right\} \\ &= P\left\{\frac{|\widehat{T}^c_{OR}(n^*_{OR}) - T^c_{OR}|}{T^c_{OR}} \leqslant \alpha\right\} = \beta. \end{aligned} \tag{E.15}$$

Since the conditional cycle overrun is assumed to follow the exponential distribution, similar to the justification of Proposition 2.3, we approximate $\widehat{T}^c_{OR}(n^*_{OR})$ by a Gaussian random variable $\mathcal{N}\left(T^c_{OR}, \frac{{T^c_{OR}}^2}{n^*_{OR}}\right)$. Therefore, we have:

$$\begin{aligned} & P\left\{\frac{|\widehat{T}^c_{OR}(n^*_{OR}) - T^c_{OR}|}{T^c_{OR}} \leqslant \alpha\right\} \\ =& P\left\{(1-\alpha)T^c_{OR} \leqslant \widehat{T}^c_{OR}(n^*_{OR}) \leqslant (1+\alpha)T^c_{OR}\right\} \\ =& P\left\{\widehat{T}^c_{OR}(n^*_{OR}) \leqslant (1+\alpha)T^c_{OR}\right\} - P\left\{\widehat{T}^c_{OR}(n^*_{OR}) \leqslant (1-\alpha)T^c_{OR}\right\} \\ \approx& \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\alpha T^c_{OR}}{\sqrt{2}\sqrt{\frac{{T^c_{OR}}^2}{n^*_{OR}}}}\right)\right) - \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{-\alpha T^c_{OR}}{\sqrt{2}\sqrt{\frac{{T^c_{OR}}^2}{n^*_{OR}}}}\right)\right) = \operatorname{erf}\left(\frac{\alpha\sqrt{n^*_{OR}}}{\sqrt{2}}\right). \end{aligned} \tag{E.16}$$

Solve for the Gaussian approximation of $n^*_{OR}(\alpha, \beta)$, we obtain:

$$n^*_{OR}(\alpha, \beta) = \left\lceil 2\left(\frac{\text{erf}^{-1}(\beta)}{\alpha}\right)^2 \right\rceil. \tag{E.17}$$

□

E.5 Justification of Proposition 6.6

For $\widehat{p}_{OR}(n^*_{p_{OR}})$ and $\widehat{k}_{OR}(n^*_{p_{OR}})$ satisfying $\frac{|\widehat{p}_{OR}(n^*_{p_{OR}})-p_{OR}|}{p_{OR}} \leqslant \alpha_1$ and $\frac{|\widehat{k}_{OR}(n^*_{k_{OR}})-k_{OR}|}{k_{OR}} \leqslant \alpha_2$, respectively, we have:

$$\begin{aligned}
&(1-\alpha_1)p_{OR} \leqslant \widehat{p}_{OR}(n^*_{p_{OR}}) \leqslant (1+\alpha_1)p_{OR}, \\
&(1-\alpha_2)k_{OR} \leqslant \widehat{k}_{OR}(n^*_{k_{OR}}) \leqslant (1+\alpha_2)k_{OR} \\
\Longleftrightarrow\ &(1-\alpha_1)(1-\alpha_2)p_{OR}k_{OR}\tau \leqslant \widehat{p}_{OR}(n^*_{p_{OR}})\widehat{k}_{OR}(n^*_{k_{OR}})\tau \leqslant (1+\alpha_1)(1+\alpha_2)p_{OR}k_{OR}\tau \\
\Longleftrightarrow\ &(1-\alpha_1-\alpha_2+O(\alpha_1\alpha_2))T_{OR} \leqslant \widehat{T}_{OR} \leqslant (1+\alpha_1+\alpha_2+O(\alpha_1\alpha_2))T_{OR} \\
\Longleftrightarrow\ &\frac{|\widehat{T}_{OR}-T_{OR}|}{T_{OR}} \leqslant \alpha_1+\alpha_2+O(\alpha_1\alpha_2).
\end{aligned} \tag{E.18}$$

Therefore, with accuracy up to $O(\alpha_1\alpha_2)$,

$$\alpha^{ind}_{OR} = \alpha_1 + \alpha_2. \tag{E.19}$$

□

E.6 Justification of Proposition 6.7

Consider the distribution $f_{OR}(t)$, its mean is $p_{OR}k_{OR}\tau$, and its variance is given by $p_{OR}k_{OR}^2\tau^2(2-p_{OR})$. Thus, the mean of $n_{p_{OR}}^*$ number of unconditional cycle overrun durations, i.e., $\widehat{T}_{OR}(n_{p_{OR}}^*)$, has the expectation $p_{OR}k_{OR}\tau$ and variance $\frac{p_{OR}k_{OR}^2\tau^2(2-p_{OR})}{n_{p_{OR}}^*}$. Based on these values, we consider approximate $\widehat{T}_{OR}(n_{p_{OR}}^*)$ by a Gaussian random variable $\mathcal{N}\left(p_{OR}k_{OR}\tau, \frac{p_{OR}k_{OR}^2\tau^2(2-p_{OR})}{n_{p_{OR}}^*}\right)$. Therefore, we have:

$$
\begin{aligned}
\beta_{OR}^{ind} &= P\left\{\frac{|\widehat{T}_{OR}(n_{p_{OR}}^*) - T_{OR}|}{T_{OR}} \leqslant \alpha_{OR}^{ind}\right\} \\
&= P\left\{\widehat{T}_{OR}(n_{p_{OR}}^*) \leqslant (1+\alpha_{OR}^{ind})T_{OR}\right\} - P\left\{\widehat{T}_{OR}(n_{p_{OR}}^*) \leqslant (1-\alpha_{OR}^{ind})T_{OR}\right\} \\
&\approx \frac{1}{2}\text{erf}\left(\frac{\alpha_{OR}^{ind}p_{OR}k_{OR}\tau}{\sqrt{\frac{p_{OR}k_{OR}^2\tau^2(2-p_{OR})}{n_{OR}^{**}}}\sqrt{2}}\right) - \frac{1}{2}\text{erf}\left(\frac{-\alpha_{OR}^{ind}p_{OR}k_{OR}\tau}{\sqrt{\frac{p_{OR}k_{OR}^2\tau^2(2-p_{OR})}{n_{OR}^{**}}}\sqrt{2}}\right) \\
&\approx \text{erf}\left((\alpha_1+\alpha_2)C\sqrt{\frac{n_{p_{OR}}^*}{2}}\right),
\end{aligned}
\tag{E.20}
$$

where $C = \sqrt{\frac{\widehat{p}_{OR}(n_{p_{OR}}^*)}{2-\widehat{p}_{OR}(n_{p_{OR}}^*)}}$. □

E.7 Justification of Proposition 6.8

Similar to the justification of Proposition 6.7, we approximate $\widehat{T}_{OR}(n)$ by a Gaussian random variable $\mathcal{N}\left(p_{OR}k_{OR}\tau, \frac{p_{OR}k_{OR}^2\tau^2(2-p_{OR})}{n}\right)$. Therefore, we have:

$$\begin{aligned}
\delta &= P\left\{\frac{|\widehat{T}_{OR}(n_{OR}^{**}) - T_{OR}|}{T_{OR}} \leqslant \gamma\right\} \\
&= P\left\{\widehat{T}_{OR}(n_{OR}^{**}) \leqslant (1+\gamma)T_{OR}\right\} - P\left\{\widehat{T}_{OR}(n_{OR}^{**}) \leqslant (1-\gamma)T_{OR}\right\} \\
&\approx \frac{1}{2}\text{erf}\left(\frac{\gamma p_{OR}k_{OR}\tau}{\sqrt{\frac{p_{OR}k_{OR}^2\tau^2(2-p_{OR})}{n_{OR}^{**}}}\sqrt{2}}\right) - \frac{1}{2}\text{erf}\left(\frac{-\gamma p_{OR}k_{OR}\tau}{\sqrt{\frac{p_{OR}k_{OR}^2\tau^2(2-p_{OR})}{n_{OR}^{**}}}\sqrt{2}}\right) \qquad (E.21) \\
&= \text{erf}\left(\sqrt{\frac{n_{OR}^{**}}{2}}\frac{\gamma}{\sqrt{\frac{2}{\widehat{p}_{OR}} - 1}}\right).
\end{aligned}$$

Solve for the Gaussian approximation of n_{OR}^{**}, we obtain

$$n_{OR}^{**}(\gamma,\delta) = \left\lceil 2\left(\frac{2-\widehat{p}_{OR}}{\widehat{p}_{OR}}\right)\left(\frac{\text{erf}(\delta)}{\gamma}\right)^2\right\rceil. \qquad (E.22)$$

BIBLIOGRAPHY

Abramowitz, M., and I. Stegun (1948), *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55, US Government printing office, Washington, D.C.

Alavian, P., Y. Eun, K. Liu, S. M. Meerkov, and L. Zhang (2018), The (α, β)-precise estimates of MTBF and MTTR: Definitions, calculations, and effect on machine efficiency and throughput evaluation in serial production lines, *URL: http://web.eecs.umich.edu/˜smm/publications/mtbf_mttr_estimates.pdf*.

Alavian, P., Y. Eun, K. Liu, S. M. Meerkov, and L. Zhang (2019), The (α, β)-precise estimates of MTBF and MTTR: Definitions, calculations, and induced effect on machine efficiency evaluation, *IFAC MIM 2019 Proceedings*, *52*(13), 1004–1009.

Alavian, P., Y. Eun, K. Liu, S. M. Meerkov, and L. Zhang (2021), The (α, β)-precise estimates of mtbf and mttr: Definition, calculation, and observation time, *IEEE Transactions on Automation Science and Engineering*, *18*(3), 1469–1477, doi:10.1109/TASE.2020.3017134.

Alavian, P., Y. Eun, K. Liu, S. M. Meerkov, and L. Zhang (2021), The (α_X, β_X)-precise estimates of production systems performance metrics, *International Journal of Production Research*, *DOI: 10.1080/00207543.2021.1886367*.

Altiok, T. (1997), *Performance Analysis of Manufacturing Systems*, Springer-Verlag, New York, NY.

Altiok, T., and B. Melamed (2010), *Simulation Modeling and Analysis with Arena*, Elsevier, Oxford, UK.

Askin, R. G., and C. R. Standridge (1993), *Modeling and Analysis of Manufacturing Systems*, vol. 29, Wiley, New York, NY.

Bai, Y., J. Tu, M. Yang, L. Zhang, and P. Denno (2021), A new aggregation algorithm for performance metric calculation in serial production lines with exponential machines: design, accuracy and robustness, *International Journal of Production Research*, *59*(13), 4072–4089.

Barlow, R., and F. Proschan (1996), *Mathematical theory of reliability*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Ben-Ammar, O., B. Bettayeb, and A. Dolgui (2020), Integrated production planning and quality control for linear production systems under uncertainties of cycle time and finished product quality, *International Journal of Production Research*, *58*(4), 1144–1160.

Boucheron, S., G. Lugosi, and O. Bousquet (2003), Concentration inequalities, in *Summer school on machine learning*, pp. 208–240, Springer, New York, NY.

Buzacott, J. A., and J. G. Shanthikumar (1993), *Stochastic Models of Manufacturing Systems*, vol. 4, Prentice Hall, Englewood Cliffs, NJ.

Casalino, A., A. M. Zanchettin, L. Piroddi, and P. Rocco (2019), Optimal scheduling of human-robot collaborative assembly operations with time petri nets, *IEEE Transactions on Automation Science and Engineering*, *18*(1), 70–84.

Chehade, A., and Z. Shi (2019), Sensor fusion via statistical hypothesis testing for prognosis and degradation analysis, *IEEE Transactions on Automation Science and Engineering*, *16*(4), 1774–1787.

Chen, W., and Z. Wang (2019), Analysis of the steady state probability distribution of a manufacturing system under the prioritised hedging point control policy, *International Journal of Production Research*, *57*(8), 2281–2303.

Chhetri, S. R., S. Faezi, A. Canedo, and M. A. A. Faruque (2019), Quilt: quality inference from living digital twins in iot-enabled manufacturing systems, in *Proceedings of the International Conference on Internet of Things Design and Implementation*, pp. 237–248.

Curry, G. L., and R. M. Feldman (2010), *Manufacturing systems modeling and analysis*, Springer Science & Business Media, New York, NY.

Eun, Y., K. Liu, and S. M. Meerkov (2021), Production systems with cycle overrun: Modeling, analysis, improvability, and bottlenecks, *URL:* `https://web.eecs.umich.edu/~smm/publications/ProductionsSystemswithCycleOverrunFullData041821.pdf`.

Fang, W., Y. Guo, W. Liao, K. Ramani, and S. Huang (2020), Big data driven jobs remaining time prediction in discrete manufacturing system: a deep learning-based approach, *International Journal of Production Research*, *58*(9), 2751–2766.

Feller, W. (2015), On the normal approximation to the binomial distribution, in *Selected Papers I*, pp. 655–665, Springer, New York, NY.

Gershwin, S. B. (1994), *Manufacturing Systems Engineering*, Prentice Hall, Englewood Cliff, NJ.

Gopalswamy, K., and R. Uzsoy (2019), A data-driven iterative refinement approach for estimating clearing functions from simulation models of production systems, *International Journal of Production Research*, *57*(19), 6013–6030.

Gubner, J. A. (2006), *Probability and random processes for electrical and computer engineers*, Cambridge University Press.

Hao, L., L. Bian, N. Gebraeel, and J. Shi (2017), Residual life prediction of multistage manufacturing processes with interaction between tool wear and product quality degradation, *IEEE Transactions on Automation Science and Engineering*, *14*(2), 1211–1224.

Hwang, G., J. Lee, J. Park, and T.-W. Chang (2017), Developing performance measurement system for internet of things and smart factory environment, *International Journal of Production Research*, *55*(9), 2590–2602.

Inman, R. R. (1999), Empirical evaluation of exponential and independence assumptions in queueing models of manufacturing systems, *Production and Operations Management*, *8*(4), 409–432.

Jerry, B. (2005), *Discrete Event System Simulation*, Pearson Education, Cranbury, NJ.

Kacar, N. B., L. Mönch, and R. Uzsoy (2016), Modeling cycle times in production planning models for wafer fabrication, *IEEE Transactions on Semiconductor Manufacturing*, *29*(2), 153–167.

Kagermann, H., J. Helbig, A. Hellinger, and W. Wahlster (2013), *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry; final report of the Industrie 4.0 Working Group*, Forschungsunion.

Khatab, A., C. Diallo, E.-H. Aghezzaf, and U. Venkatadri (2019), Integrated production quality and condition-based maintenance optimisation for a stochastically deteriorating manufacturing system, *International Journal of Production Research*, *57*(8), 2480–2497.

Kontar, R., S. Zhou, J. Horst, and B. A (2017), Estimation and monitoring of key performance indicators of manufacturing systems using the multi-output gaussian process, *International Journal of Production Research*, *55*(8), 2304–2319.

Kullback, S. (1997), *Information theory and statistics*, Courier Corporation, Chelmsford, MA.

Kuo, C.-J., C.-F. Chien, and J.-D. Chen (2011), Manufacturing intelligence to exploit the value of production and tool data to reduce cycle time, *IEEE Transactions on Automation Science and Engineering*, *8*(1), 103–111.

Larco, J. A., R. De Koster, K. J. Roodbergen, and J. Dul (2017), Managing warehouse efficiency and worker discomfort through enhanced storage assignment decisions, *International Journal of Production Research*, *55*(21), 6407–6422.

Law, A. M., and W. D. Kelton (1991), *Simulation Modeling and Analysis*, vol. 2, McGraw-Hill, New York, NY.

Li, J., and S. M. Meerkov (2009), *Production systems engineering*, Springer Science & Business Media, New York, NY.

Liao, Y., F. Deschamps, E. de Freitas Rocha Loures, and L. F. P. Ramos (2017), Past, present and future of industry 4.0 - a systematic literature review and research agenda proposal, *International Journal of Production Research*, *55*(12), 3609–3629, doi:10.1080/00207543.2017.1308576.

Lin, J. (1991), Divergence measures based on the shannon entropy, *IEEE Transactions on Information theory*, *37*(1), 145–151.

Lin, Z., A. Matta, and J. G. Shanthikumar (2019), Combining simulation experiments and analytical models with area-based accuracy for performance evaluation of manufacturing systems, *IISE Transactions*, *51*(3), 266–283.

Marsaglia, G. (2006), Ratios of normal variables, *Journal of Statistical Software*, *16*(4), 1–10.

Meerkov, S. M., and C.-B. Yan (2016), Production lead time in serial lines: Evaluation, analysis, and control, *IEEE Transactions on Automation Science and Engineering*, *13*(2), 663–675.

Meerkov, S. M., and L. Zhang (2010), Product quality inspection in bernoulli lines: analysis, bottlenecks, and design, *International Journal of Production Research*, *48*(16), 4745–4766.

Millstein, M. A., and J. S. Martinich (2014), Takt time grouping: implementing kanban-flow manufacturing in an unbalanced, high variation cycle-time process with moving constraints, *International Journal of Production Research*, *52*(23), 6863–6877.

Morrison, J. R., and D. P. Martin (2007), Practical extensions to cycle time approximations for the $g/g/m$-queue with applications, *IEEE Transactions on Automation Science and Engineering*, *4*(4), 523–532.

Muchiri, P. N., L. Pintelon, H. Martin, and P. Chemweno (2014), Modelling maintenance effects on manufacturing equipment performance: results from simulation analysis, *International Journal of Production Research*, *52*(11), 3287–3302.

Müller, M., M. Lehmann, and H. Kuhn (2020), Measuring sequence stability in automotive production lines, *International Journal of Production Research*, dOI: 10.1080/00207543.2020.1790685.

Nadarajah, S., and S. Kotz (2008), The cycle time distribution, *International Journal of Production Research*, *46*(11), 3133–3141.

Papadopolous, H. T., C. Heavey, and J. Browne (1993), *Queueing Theory in Manufacturing Systems Analysis and Design*, Springer Science & Business Media, New York, NY.

Papadopoulos, H. T., M. E. J. O'Kelly, M. J. Vidalis, and D. Spinellis (2009), *Analysis and Design of Discrete Part Production Lines*, Springer, New York, NY.

Perros, H. G. (1994), *Queueing networks with blocking*, Oxford University Press, Inc., Oxford, UK.

Resnick, S. (2019), *A probability path*, Springer, New York, NY.

Roshani, A., M. Paolucci, D. Giglio, and F. Tonelli (2020), A hybrid adaptive variable neighbourhood search approach for multi-sided assembly line balancing problem to minimise the cycle time, *International Journal of Production Research*, p. DOI: 10.1080/00207543.2020.1749958.

Saez, M., F. P. Maturana, K. Barton, and D. M. Tilbury (2018), Real-time manufacturing machine and system performance monitoring using internet of things, *IEEE Transactions on Automation Science and Engineering*, *15*(4), 1735–1748.

Schneckenreither, M., S. Haeussler, and C. Gerhold (2020), Order release planning with predictive lead times: a machine learning approach, *International Journal of Production Research*, DOI: 10.1080/00207543.2020.1859634.

Touzani, H., H. Hadj-Abdelkader, N. Seguy, and S. Bouchafa (2021), Multi-robot task sequencing & automatic path planning for cycle time optimization: Application for car production line, *IEEE Robotics and Automation Letters*, *6*(2), 1335–1342.

Tsagris, M., C. Beneki, and H. Hassani (2014), On the folded normal distribution, *Mathematics*, *2*(1), 12–28, Multidisciplinary Digital Publishing Institute.

Viswanadham, N., and Y. Narahari (1992), *Performance Modeling of Automated Manufacturing Systems*, Prentice Hall, Englewood Cliff, NJ.

Williams, E. J. (1994), Downtime data – Its collection, analysis, and importance, in *Proceedings of the 26th conference on Winter simulation*, pp. 1040–1043.

Yu, C., and A. Matta (2016), A statistical framework of data-driven bottleneck identification in manufacturing systems, *International Journal of Production Research*, *54*(21), 6317–6332.

Zhou, P., S.-W. Lu, and T. Chai (2014), Data-driven soft-sensor modeling for product quality estimation using case-based reasoning and fuzzy-similarity rough sets, *IEEE Transactions on Automation Science and Engineering*, *11*(4), 992–1003.