Cuplikan Buku
[Terkait dengan *Copyright Agreement*]

Reliable Shewhart-type Control Charts for Multivariate Process Variability

Oleh:
Maman Djauhari
Zunna'aim Zolkeply
Dyah Erny Herwindiati

Publising Lambert

Maman Djauhari
Zunna'aim Zolkeply
Dyah Herwindiati

Reliable Shewhart-type Control Charts for Multivariate Process Variability

LAP LAMBERT Academic Publishing

Maman Djauhari
Zunna'aim Zolkeply
Dyah Herwindiati

Reliable Shewhart-type Control Charts for Multivariate Process Variability

LAP LAMBERT Academic Publishing

Imprint

Cover image: www.ingimage.com

Publisher:
LAP LAMBERT Academic Publishing
is a trademark of
International Book Market Service Ltd., member of OmniScriptum Publishing Group
17 Meldrum Street, Beau Bassin 71504, Mauritius

Printed at: see last page
ISBN: 978-613-9-58609-7

RELIABLE SHEWHART-TYPE CONTROL CHARTS FOR MULTIVARIATE PROCESS VARIABILITY

By:
Maman Abdurachman Djauhari,
Zunna'aim Zolkepley and Dyah Erny Herwindiati

FOREWORD FROM THE PRESIDENT OF MALAYSIAN ACADEMY OF MATHEMATICAL SCIENTISTS

The experience of advanced nations shows that the industrial sector especially in manufacturing is one of the important factors to generate economic growth of a nation. The role of the industrial sector to provide opportunities for employment, inventing new technologies and sustain economic growth is of great importance. The driving force behind the progress of the industrial sector needless to say is the knowledge-workers. A nation needs to develop such human capital in order for it to continue progress and advance to more elevated level with sound economic activities that will lead to generation of higher income among its populace. It is one of the pillars that will support and maintain social stability of a nation. It is due to this that the industrial sector ought to remain vibrant and sustain its strength for as long as it is possible.

One of the factors that are essential in maintaining strength and always ready to face up to any challenges as time changes that the main players in industry should not neglect is the quality of the manpower that shoulder the task of maintaining quality of its production processes. Exposure must be given to the management of industries and their workforce to the tools that are needed to do so. They should be knowledgeable in the theoretical and practical knowledge of the task they are shouldering. At the management level at least visionary personnel are needed who have the ability to forecast the future needs of the communities and design plans and programs to meet the forthcoming demands of the consumers. In this respect to have the deep knowledge, skill and ability to measure the quality of production process is indeed essential.

In this book the authors highlight the essential ingredients in multivariate statistical process control. With ample background to the development of measuring the quality of production process since early twenties of the last century to the present the authors have proposed a more refined method to do so. Led by Maman Djauhari the authors have successfully present their thoughts in clarity which I feel is easily understood by practitioners in this field. It offers an understanding on the current practice in multivariate statistical process control and its limitations, proposing the

TABLE OF CONTENTS

CHAPTER 1
INTRODUCTION

In statistical process control (SPC), control chart is the main source of information about the history of process performance. This information leads to a decision making on the process quality. Therefore, control charts must be reliable. In practice, their reliability is determined by,

(i) The appropriateness of statistic used to construct the control chart,

(ii) The control limits which reflect the desired probability of false alarm (PFA).

It is worth noting that the word "reliable" is an adjective. It means "consistently good in quality or performance." Therefore, a reliable control chart is the one which is "able to be trusted."

This book is focused on multivariate statistical process control (MSPC), a branch of SPC dealing with controling/monitoring simultaneously several correlated variables (also called quality characteristics). And more specifically, it is focused on multivaiate process variability (MPV) monitoring. As remarked in the literature such as Montgomery (2001, 2005, 2009), all quality professionals know that MPV is as important as process target to be monitored. The techniques available to conduct this monitoring operation will help us to detect the special causes of MPV. When a process is statistically under control, then only common causes affect MPV.

In the last few years MPV monitoring operation has been thoroughly studied. And nowadays, many techniques and tools for this operation are available in the literature. However, according to Ries and Rato (2013), GV-based control charts are still the most adopted control charts to do this job. For example, recently Carlos Garcia-Diaz (2007) has developed a control chart based on the so-called effective variance (EV). A special feature of this chart is that it can be used to compare the variability of two processes with different number of quality characteristics. And we know that EV is a variant of generalized variance (GV). Due to its importance, we start our discussion on the *GV*-chart.

GV-chart is easy to construct and to interprete. The geometrical interpretation and distributional behavior of GV are interesting subject of discussion in the literature since the last five decades. See, for exampe, Anderson (1966, 1984) for an early development of the geometrical interpretation, Mason et al. (2009) the geometrical interpretation in MSPC, and Djauhari (2009) for a distributional behavior. This is perhaps the reason why, despite of serious limitations of GV as a measure of MPV, it becomes the most popular and widely used in practice.

Alt and Smith (1988), see also Montgomery (2001, 2005, 2009) among others, have specified the limitations of GV. This motivates Djauhari (2007) to questioning the appropriateness of GV as a measure of MPV. Then, in that paper he introduces the so-called vector variance (VV) as another measure and shows that VV and GV are complementary. The advantage of *VV*-chart in MPV monitoring compared to *GV*-chart, is demonstrated in Djauhari et al. (2008) and a recommendation to use them simultaneously is given in Djauhari and Mohamad (2010). This leads us to put *VV*-chart as the second topic of discussion in this book.

The problem reveals when we deal with small sample size. A preliminary study in Djauhari et al. (2016), involving samples of small, moderate and large (but not sufficiently large) size, shows that the control limits of both *GV*-chart and *VV*-chart have nothing to do with the probability of false alarm (PFA); they do not reflect the PFA. This means that the presence of an out-of-control (OOC) signal cannot probabilistically be trusted. Its reliability is unknown. Through simulation experiments, in this book we show that the results of this preliminary study are justified. Therefore, we expand this study to a more general sample size with general number of quality characteristics. Furthermore, to overcome the non-reliability problem, a reliability constant is introduced and simulation experiments were conducted to find its value. This will leads us to the construction of a reliable *GV*-chart and a reliable *VV*-chart. Finally, to facilitate the practitioners with these reliable charts, tables of reliability constant are provided and industrial examples are presented.

The remaining chapters are organized as follows. Chapter 2 is devoted to recall the most popular and important variability measures in univariate case and multivariate case as well. It is followed by Chapter 3 which focuses on classical *GV*-chart. "Classical" means that the control limits for small sample size are determined heuristically while normal approximation is used for large sample size. Later on, a reliable *GV*-chart for small sample size is discussed in Chapter 4. In this chapter, tables of reliability constant for *GV*-chart are provided for some selected PFAs. Then, an industrial example is presented to illustrate the need for reliable control chart. Our discussion is continued with *VV*-chart in Chapter 5. The book ends in Chapter 6 with a reliable *VV*-chart for small sample size, tables of reliability constant for *VV*-chart for some selected PFAs, and an industrial example to illustrate the advantage of this chart.

CHAPTER 2
MULTIVARIATE PROCESS VARIABILITY MEASURES

Process is anywhere and everywhere. Statistically, its behavior is represented by the so-called probabilistic distribution that governs the process. In this language, process target and process variability are represented by the so-called measure of central tendency and measure of dispersion, respectively. The central tendency refers to a location at which the distribution is concentrated. Meanwhile, the dispersion refers to the way all possible values of the variable(s) under study are scattered. Since the concern of this book is on process variability, in this chapter we recall some important measures of variability in multivariate process. Just for the sake of continuity, some measures of univariate process variability will also be recalled but not discussed. By univariate process we mean the process where only one quality characteristic is to be monitored. If there are more than one quality characteristic, and their correlations are taken into account, the process is called a multivariate process.

In univariate process, range, variance and standard deviation are the most widely used variability measures in SPC. There are also other measures such as Interquartile range, Luceno's measure, and quantile-based measure. Interquartile range is the range between the third quartile Q_3 and the first Q_1. What are the last two measures? Luceno's measure is defined as

$$L = \frac{1}{n}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{n}\left|X_i - \bar{X}\right| \text{ where } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

where X_1, X_2, ..., X_n is a sample of size n. Meanwhile, quantile-based measure is

$$\frac{x_{0.99865} - x_{0.00135}}{6}$$

where x_α is the $(1-\alpha)$-th quantile. These last two measures are derived from the capability indices C_{pc} and $C_p(q)$, respectively. A good discussion of these measures can be found in Montgomery (2001, p. 364-365).

In multivariate process, the variability of the process is difficult to measure. Numerically, it is represented in the form of covariance matrix. The difficulty appears because there is no single scalar measure that can be used to represent the complex covariance structure contained in this matrix. However, the following two measures are the important ones that usually used in multivariate statistical process control (MSPC); generalized variance (GV) and vector variance (VV). Other measures are also available in the literature of multivariate analysis but not very much appreciated in MSPC. We can say, for example, total variance (TV), square root of generalized variance (SRGV), relative generalized variance (RGV) and minimum volume ellipsoid (MVE). Some authors such as Alt and Smith (1988) and Montgomery (2001, 2005) have discussed the application of SRGV in MSPC. See also Djauhari (2005) for an improved version of SRGV. However, in practice, SRGV is dominated by GV. Meanwhile, an application of RGV can be found in Tang and Barnett (1996). Regarding MVE, it is usually empoyed in the literature of robust statistics. See, for example, Rousseeuw (1985), and Grambow and Stromberg (1998) for the details.

Another example is effective variance (EV). It is the geometric mean of all eigenvalues of Σ and defined as $\sqrt[p]{|\Sigma|}$. See Serfling (1980) and Pena and Rodriguez (2005) for the definition and Carlos Garcia-Diaz (2007) for its application in MSPC. This measure is usually used to compare two processes, in terms of their variability, when these processes involve different number of quality characteristics. It is worth noting that EV, is a variant of GV. Therefore, it inherits the property of GV.

Due to the limitation of EV, caused by the limitation of GV, in this book we introduce another measure based on VV called "effective vector variance (EVV)." We define EVV $= \frac{1}{p}Tr\left(\Sigma^2\right)$. It is the arithmatics mean of the squared eigenvalues of Σ. Here, p is the number of quality characteristics involved in the process monitoring. Like EV, EVV can also be used to compare two processes involving different number of quality characteristics.

In the rest of the book the discussion will be focused on GV and VV. We start with GV, the most popular and widely used measure in MSPC. Its advantages as well as its weaknesses will be discussed. It is to handle its weaknesses that VV is introduced as an alternative variability measure. We show that a simultaneous use of both GV and VV will be advantageous.

2.1. Generalized variance

There is no doubt that GV has a special role in monitoring covariance structure stability in a wide range of scientific investigations; from soft sciences to hard sciences, and from service industry to manufacturing industry. We can mention, for example, Beamon and Ware (1998) who employ GV in supply chain management, Edelman and Rao (2005) in theoretical physics, and Hubert et al. (2008) in astronomical data. We can also mention Wood (1994), Roes and Dorr (1997), and Sulek (2004) for its application in service industry, Sellick (1993), Shahian et al. (1996), Hanslik et al. (2001), and Woodall (2006) in health care industry, Ye et al. (2001) and Kruegel et al. (2005) in information industry, Florac et al. (2000) and Jakolte and Saxena (2002) in software industry, Ragea (2003) and Da Costa et al. (2005) in financial industry, and Woodall and Montgomery (1999), Djauhari (2005), Sullivan et al. (2007), and Mason et al. (2009) in manufacturing industry. Many more applications can easily be found in the literature showing the importance role of GV in real application.

In the field of MSPC, all quality professionals know that process variability is very important to be monitored. It is as important as monitoring the process target. They understand that the former is more difficult than the latter. This is most probably caused by the fact that a scalar measure, and even several scalar measures used simultaneously, cannot be used to explain the whole covariance structure summarized in Σ. Take GV as an example. It is the determinant Σ. Therefore, according to the property of determinant operator, two different processes could have the same GV. In other words, according to GV, two different processes could have the same variability. Thus, when *GV*-chart is used to monitor the process variability, the chart is only a necessary chart. In other words if the process is in-control, then *GV*-chart is in-control but not vice versa. Furthermore, its sampling distribution tends very slowly to normality. Thus, it is not apt for practical purpose because in real application the sub-group size (i.e. the sample size) n is very limited.

One of the most important advantages of GV lies in its sampling distribution. As we will see in Sub-section 2.1.4, for $p > 2$, its distribution can be approximated by a very familiar distribution, say, normal distribution when n is sufficiently large. See also Muirhead (1982) and Anderson (1984) for the proof. However, when n is small, normal approximation is no longer appropriate. In this case, as can be seen in the literature such as Montgomery (2005, 2009), the current practice of process variability monitoring operation is by using *GV*-chart based on heuristic approach. The control limits are derived heuristically from the property that "*most of the probability distribution of* $|S|$ *is contained in the interval* $E(|S|) \pm 3\sqrt{Var(|S|)}$."

In the next chapter, we show that this property will lead to a non-reliable control chart. This is a serious weakness of *GV*-chart for small *n*. For further discussions on the advantages and weaknesses of GV, the readers are suggested to consult Alt and Smith (1988), Montgomery (2005) and Djauhari (2005).

Despite its weaknesses, due to the above desirable property, in the next paragraph we study the sampling distribution of GV. Before doing so, we first recall the classical asymptotic distribution. Later on, the parameter estimates issued from several independent random samples will also be derived. These estimates will provide the basic tools in process variability monitoring.

In the next sub-section, we start our presentation with geometrical interpretation of GV followed consecutively by its distributional behavior, delta method to derive the distribution of a function of random variables, and finally the asymptotic sampling distribution of GV.

2.1.1. Geometrical interpretation

In this book, unless otherwise, a random sample $X_1, X_2, \ldots, X_n$ of size *n* will always be considered drawn from $N_p(\mu, \Sigma)$; a *p*-variate normal distribution with mean vector μ and covariance matrix Σ positive definite. We assume $n > p$. The mean vector and covariance matrix of this sample are,

$$\bar{X} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_p \end{pmatrix} \text{ and } S = \begin{pmatrix} s_{11} & s_{12} & \cdot & \cdot & \cdot & s_{1p} \\ s_{21} & s_{22} & \cdot & \cdot & \cdot & s_{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ s_{p1} & s_{p2} & \cdot & \cdot & \cdot & s_{pp} \end{pmatrix}.$$

where $\bar{x}_i$ is the sample mean of the *i*-th variable and s_{ij} is the sample covariance of

CHAPTER 3
GENERALIZED VARIANCE CHART

In the rest of the book, process variability monitoring will always be based on m independent samples of the same size n, each of which is drawn from a p-variate normal distribution with positive definite covariance matrix Σ. And we focus in this chapter on monitoring operation using *GV*-chart discussd in Chapter 2. For this purpose, we start our discussion with the classical technique of the *GV*-chart construction for both large and small sample size n. Then, we show that in those two cases the same control chart is employed. This is presented in Section 3.1.

After showing that the control limits of *GV*-chart are not unbiased, in Section 3.2 a new version of *GV*-chart, called improved *GV*-chart or simply denoted by *IGV*-chart, is presented. It is followed by a sensitivity analysis in terms of average run length (ARL). Later on, we show that *IGV*-chart has better ARL than *GV*-chart.

To illustrate the advantages of *IGV*-chart, an industrial example will be studied, and to make the readers comfortable with the practical implementation of this chart, in the last section we address a message for practitioners.

3.1. Classical technique

Let S_i be the covariance matrix of the i-th sample; $i = 1, 2, \ldots, m$, and $\overline{S}$ be their average. *GV*-chart is a control chart constructed by plotting $|S_i|$ and the control limits in the same chart. Theoretically, for the i-th sample, the control limits are derived from the distribution of $|S_i|$. In what folows, these control limits are studied.

We know from (2.6) that,

$$\frac{1}{\sqrt{b_2}}(|S_i| - b_1 |\Sigma|) \xrightarrow{d} N\left(0, |\Sigma|^2\right). \tag{3.1}$$

Therefore, for n sufficiently large, (3.1) allows us to approximate the distribution of $|S_i|$ by normal distribution with mean $b_1|\Sigma|$ and variance $b_2|\Sigma|^2$. In practice, this approximation is far from the limit because the convergence is very slow. This indicates that the practicality of the *GV*-chart for small n will differ considerably from the onę where n is large.

3.1.1. Case of large sample

If Σ is unknown, the parameters mean and variance in (3.1) must be estimated. Based on the i-th sample alone, in Chapter 2 we show that $\frac{|S_i|}{b_1}$ and $\frac{|S_i|^2}{(b_1^2+b_2)}$ are unbiased estimates of $|\Sigma|$ and $|\Sigma|^2$, respectively. What wil happen with these estimates if all m samples are involved? Since we deal with m independent samples, in classical technique such as presented in the literature, $|\Sigma|$ is estimated by $\frac{|\bar{S}|}{b_1}$ and $|\Sigma|^2$ by $\frac{|\bar{S}|^2}{(b_1^2+b_2)}$. Then, these estimates are used directly to define the control limits.

Specifically, for probability of false alarm (PFA) $\alpha = 0.0027$, this technique defines the upper control limit (UCL), central line (CL), and lower control limit (LCL) as follows,

$$\begin{aligned} \text{UCL} &= \frac{|\bar{S}|}{b_1}(b_1 + 3\sqrt{b_2}) \\ \text{CL} &= |\bar{S}| \\ \text{LCL} &= \max\left\{0, \frac{|\bar{S}|}{b_1}(b_1 - 3\sqrt{b_2})\right\}. \end{aligned} \tag{3.2}$$

The *GV*-chart is then constructed by plotting, in the same diagram, the value of GV for each sample and the control limits (3.2).

This chart indicates that there is no out-of-control (OOC) signal occurs during the production process.

Before we leave this section, it is worth noting that,

1. The form of *GV*-chart is the same whether we consider $n = 5$ as large or small. However, if we consider $n = 5$ as large, PFA = 0.0027 is far from reality. On the other hand, if $n = 5$ is considered as small, than we do not know the PFA.
2. The computation process and the drawing are realized using Microsoft Excel. This is to illustrate the simplicity of *GV*-chart construction.

3.2. Recent development

In the previous section, the control limits of *GV*-chart in (3.2) are determined by using these statistics $\frac{|\bar{S}|}{b_1}$ and $\frac{|\bar{S}|^2}{(b_1^2 + b_2)}$ as the estimate of $|\Sigma|$ and of $|\Sigma|^2$. In the next sub-section we show that these estimates are not unbiased. Later on, we present the unbiased ones.

3.2.1. IGV-chart construction

In one sample case, see Montgomery (2001, 2005, 2009), we have $S \sim W_p(\Sigma,(n-1))$ which leads to the distribution of $|\bar{S}|$ in (2.1). Since the m samples are independent, then $\bar{S} \sim W_p(\Sigma, m(n-1))$ and analogous to (2.1), we have the following distribution of $|\bar{S}|$,

$$|\bar{S}| \sim \frac{|\Sigma|}{(m(n-1))^p} \prod_{k=1}^{p} \chi^2_{m(n-1)-k+1} \tag{3.3}$$

where $\chi^2_{m(n-1)}, \chi^2_{m(n-1)-1}, \ldots,$ and $\chi^2_{m(n-1)-p+1}$ are independent. This distribution leads us to the following unbiased estimates of $|\Sigma|$ and $|\Sigma|^2$,

$$|\bar{S}|\frac{b_1}{b_3} \text{ and } \frac{|\bar{S}|^2}{(b_3^2 + b_4)} \tag{3.4}$$

where

In manufacturing industry, data collection is expensive and time consuming. Therefore, to monitor the quality of production process, the use of small sample is strongly required. Unfortunately, the current techniques of multivariate process variability monitoring for small sample are developed based on heuristic approach. As a consequence, their reliability is unknown. This book introduces a reliable version of GV-Chart (the most adopted technique) and a reliable version of VV-Chart. Since both charts are complementary to each other, their simultaneous use is recommended. Furthermore, for practical purposes, the book is completed with statistical tables of reliability constant for some selected sample sizes and selected number of quality characteristics. Real industrial applications are also provided to illustrate the advantages of the two reliable charts.

Maman A. Djauhari was born in Dayeuh Garut, Tanah Pasundan, Indonesia, in 1948. He got his Doctorandus in Mathematics from Institut Teknologi Bandung (ITB), Indonesia, in 1974. His Doctorate in Pure and Applied Mathematics was obtained from Université Montpellier 2, France, in 1979. He was Dean and Chairman of Professors Council at ITB.

978-613-9-58609-7

BUKU REFERENSI – DILENGKAPI DENGAN TESTIMONI DI BAGIAN AKHIR

BIDANG B

JUDUL BUKU :

Reliable Shewhart-type Control Charts for Multivariate Process Variability

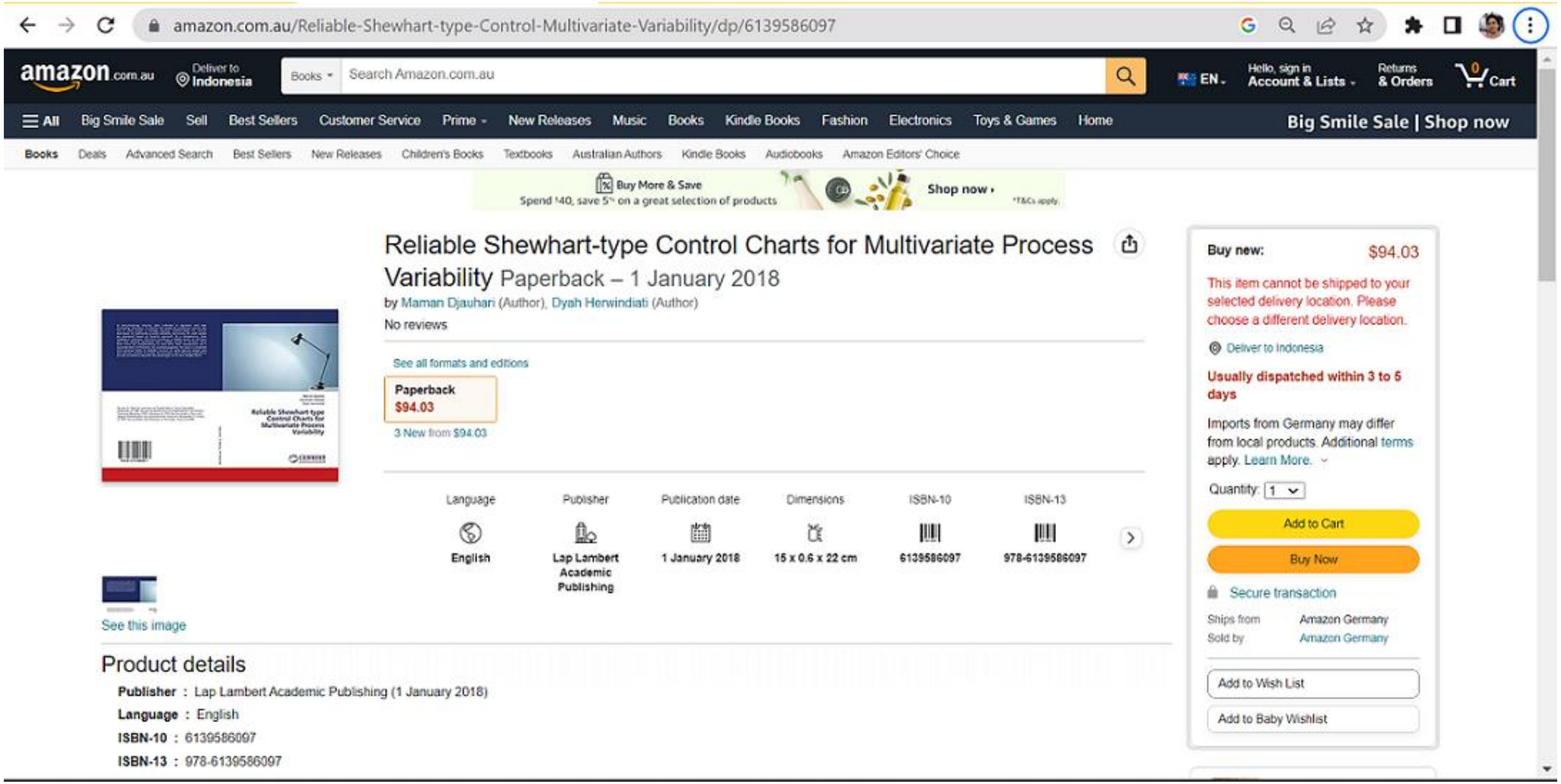
- **Publisher :** Lap Lambert Academic Publishing (1 January 2018)
- **Language :** English
- **ISBN-10 :** 6139586097
- **ISBN-13 :** 978-6139586097

Link buku:

https://www.amazon.com.au/Reliable-Shewhart-type-Control-Multivariate-Variability/dp/6139586097

Link buku:

https://www.amazon.com.au/Reliable-Shewhart-type-Control-Multivariate-Variability/dp/6139586097



UNTAR
FAKULTAS
TEKNOLOGI INFORMASI

SURAT TUGAS

Nomor : 059-PD.I/256/FTI-UNTAR/III/2018

PENUGASAN PENULISAN JURNAL NASIONAL

1.	Pejabat yang berwenang memberi tugas	: Pudek I Fakultas Teknologi Informasi
2.	Nama yang ditugaskan	: Prof. Dr. Dyah Erny Herwindiati
3.	Pangkat dan Golongan	: Guru Besar
4.	Jabatan	: Dosen Tetap
5.	Maksud dan Tujuan	: Sebagai Penulis **Buku Internasional Penerbit Lambert Academic Publishing**
7.	ISBN JURNAL	: 978-613-9-58609-7
8.	Nama Buku	: Lambert Academic Publishing
9.	Judul Buku	: Reliable Shewhart - Type Control Charts for Multivariate Process Variability

Jakarta, 29 Maret 2018
Pudek I

Ir. Jeanny Pragantha, M.Eng

Tembusan :

1. Ketua Program Studi
2. Kabag. Tata Usaha
3. Kasubag Keuangan, Personalia