Journal of Manufacturing Systems 36 (2015) 263–273

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical paper

On the effect of measurement errors in regression-adjusted monitoring of multistage manufacturing processes

Guoliang Ding[a], Li Zeng[b,*]

[a] Department of Precision Machinery and Instrumentation, University of Science and Technology of China, Hefei, Anhui 230026, China
[b] Department of Industrial and Manufacturing Systems Engineering, The University of Texas at Arlington, 500 West First Street, P.O. Box 19017, Arlington, TX 76019, USA

ARTICLE INFO

Article history:
Received 15 August 2012
Received in revised form
24 November 2013
Accepted 24 June 2014
Available online 22 July 2014

Keywords:
Measurement errors
Multistage manufacturing process (MMP)
Ordinary least squares (OLS)
Regression adjustment
Total least squares (TLS)

ABSTRACT

A critical challenge in multistage process monitoring is the complex relationships between quality characteristics at different stages. A popular method to deal with this problem is regression adjustment in which each quality characteristic is regressed on its preceding quality characteristics and the resulting residual is monitored to detect changes in local variations. However, the performance of this method depends on the accuracy of the regression coefficient estimation. One source of the estimation errors is measurement errors which commonly exist in practice. To provide guidance on the use of regression-adjusted monitoring methods, this study investigates the effect of measurement errors on the bias of regression estimation theoretically and numerically. Two estimators, the ordinary least squares (OLS) estimator and the total least squares (TLS) estimator, are compared, and insights regarding their performance are obtained.

1. Introduction

Multistage manufacturing processes (MMPs) are becoming increasingly common in today's manufacturing arena [1]. Fig. 1 shows a typical example of such processes which consists of 11 stations and is capable of producing five different types of motor reducers. A defining feature of MMPs is that the outgoing product quality at each single station is determined not only by various *local* disturbances at that station such as thermal error, cutting-force induced error, and machine geometric error, but also by the *propagated* variations from upstream stations such as the datum error due to preceding cutting operations. The general model of multistage processes in quality monitoring is shown in Fig. 2, where each node represents a quality characteristic (QC) measured at a certain stage. Due to the variation propagation in the process, these QCs bear complex relationships. This poses significant challenges for process monitoring because the conventional statistical monitoring methods are not able to differentiate local and propagated variations and thus considerable amounts of false alarms could be generated, i.e., the monitoring scheme may mistake the propagated variation as local variation and then generate an alarm that is due to other stages.

Many efforts have been made to conquer this problem by making use of either the physical models of the processes [e.g., 2–5] or statistical analysis of historical data [e.g., 6,7]. A popular data-driven method for monitoring correlated QCs is the regression adjustment method [e.g., 8,9,10,11,12]. Basically, this method monitors the residual, $Z_j = Q_j - \hat{Q}_j$, $j = 1, \ldots, q$, resulted when QC j is regressed on all its preceding QCs instead of monitoring Q_j itself. Since the propagated variation, represented by the predictor $\hat{Q}_j$, is removed, the residual will only contain the information on the local variation of QC j, and thus if Z_j is out of control, it means directly that some local faults happened. The idea of regression adjustment has been widely accepted as a simple and effective way to deal with multistage quality control problems and become the basis for many further studies.

However, the performance of regression-adjusted monitoring depends closely on the accuracy of coefficient estimation in the regression between each QC and its preceding QCs. There are two sources of estimation errors: sampling uncertainty due to limited sample size and measurement errors in data collection. Shu et al. [13–15] conduct a systematic study of the effects of the first type of estimation error on the performance of regression-adjusted monitoring. It is found that the estimation error will decrease the in-control average run length and increase the out-of-control

* Corresponding author. Tel.: +1 817 272 3150; fax: +1 817 272 3406.
E-mail address: lzeng@uta.edu (L. Zeng).

http://dx.doi.org/10.1016/j.jmsy.2014.06.013

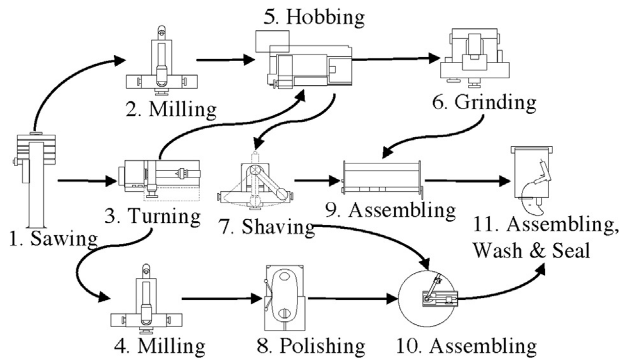

Fig. 1. An example of multistage processes.

average run length. Zeng and Zhou [16] consider the effects of the second type of estimation error in the case of large samples. They point out that the existence of measurement errors will cause inaccurate estimation of the coefficients in the regression models which eventually leads to increased false alarms and miss detections in regression-adjusted monitoring. It is worth mentioning that the effects of measurement errors in regression-adjusted monitoring have also been investigated by Li and Huang [17]. However, their study assumes that a preliminary dataset not subject to measurement errors is available from which accurate estimates of the coefficients can be obtained. So their discussion is not on the effect of estimation errors caused by measurement errors in regression-adjusted monitoring.

Since measurement errors commonly exist in manufacturing processes [18,19], it will be very useful to investigate the effects of such errors in regression-adjusted monitoring. Essentially, this means to study the estimation error caused by measurement errors. In the study of Zeng and Zhou [16], not much detail on this is provided as the focus of that study is the effects of such estimation errors on the performance of monitoring. To fill the gap, our study concentrates on the effects of measurement errors in coefficient estimation. Moreover, we compare the performance of two popular estimation methods, ordinary least squares (OLS) and total least squares (TLS), in the presence of measurement errors. The results will provide intuitive insights on regression-adjusted process monitoring as well as guidelines on its use in practice. It deserves to point out that to be useful to real-world multistage processes, this study uses an engineering model (a linear state space model) rather than a general regression model as used in many other studies [e.g., 13–15] to characterize the variation flow in the process, which is popular in multistage process research [e.g., 20]. In addition, like in [16], we assume large samples are available; in other words, we are studying the large-sample properties of the estimators. This is often the case in today's manufacturing processes due to the advancement of information/sensing technologies.

Specifically, our study aims to address two concerns on the effect of measurement errors: (1) What is the effect of measurement errors on the coefficient estimation and how is this effect affected by important characteristics of multistage processes? These characteristics include the relationships between QCs at different stages and magnitudes of local variation sources and measurement errors at each stage. (2) What are the advantages/disadvantages of the TLS estimator compared to the OLS estimator, and under what conditions it can be used to replace the OLS estimator in order to alleviate the effect of measurement errors? Both theoretical analysis and numerical study have been done to answer these questions.

The remainder of the paper is organized as follows. Section 2 will give some background information, including the process model, conventional procedure of regression-adjusted monitoring, review of estimation methods in the presence of measurement errors, and basics of the TLS method. Theoretical results on the OLS estimator and the TLS estimator will be presented in Section 3. Section 4 gives the results of a numerical study. Section 5 summarizes our findings and discusses implications of the findings on multistage process monitoring.

2. Background and basics

2.1. Process model in the presence of measurement errors

A linear state space model will be used in this study. Assume there are q QCs distributed at n stages in a general multistage process, as shown in Fig. 2. For $j = 1, \ldots, q$, define $\mathcal{P}_j = \{1, 2, \ldots, p\}$ as the set of QCs in preceding stages of QC j. For example, in Fig. 2, $\mathcal{P}_5 = \{1, 2, 3\}$, while $\mathcal{P}_q$ includes all the QCs except q. In this study, we assume that QC j could only be influenced by QCs in $\mathcal{P}_j$. Also, let U_j be the local variation source of QC j. The local variation source represents the quantities which are related to a specific QC and are often not directly observable. They have different physical meanings in different processes. For simplicity, we assume every QC in the process has a different local variation source. A linear model is assumed for Q_j and Q_i, $i \in \mathcal{P}j$

$$Q_j = \sum_{i \in P_j} \beta_{ij} Q_i + U_j \tag{1}$$

where β_{ij} is the coefficient of Q_i in the model of Q_j. In the presence of measurement errors, the observed quantities are $(Y, X_1, \ldots, X_p)$ which satisfy

$$\begin{aligned} Y &= Q_j + e \\ X_i &= Q_i + \varepsilon_i, \quad i = 1, \ldots, p \end{aligned} \tag{2}$$

where Y is the observation of QC j, X_i is the observation of QC i, e is the measurement error of QC j and ε_i is that of QC i. As in many studies on multistage process monitoring [e.g., 1], we assume that all the local variation sources and measurement errors follow normal distribution, have zero mean and are independent of each other.

2.2. Regression-adjusted monitoring

In regression-adjusted monitoring, QC j, $j = 1, \ldots, q$, is monitored using a univariate control chart to detect changes in its local variation, i.e., U_j. Following a standard procedure in SPC practice, the monitoring scheme includes two steps: In Phase I analysis, the relationship of QC j and its preceding QCs is estimated by the ordinary least squares method

$$\hat{\boldsymbol{\beta}}^{\mathrm{OLS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y \tag{3}$$

where $\hat{\boldsymbol{\beta}}^{\mathrm{OLS}} = \left[\hat{\beta}_{1j}, \hat{\beta}_{2j}, \ldots, \hat{\beta}_{pj}\right]'$, and $\mathbf{X} = [X_1,\ X_2,\ \ldots,\ X_p]'$. The resulting residual is

$$Z = Y - \mathbf{X}\hat{\boldsymbol{\beta}}^{\mathrm{OLS}}$$

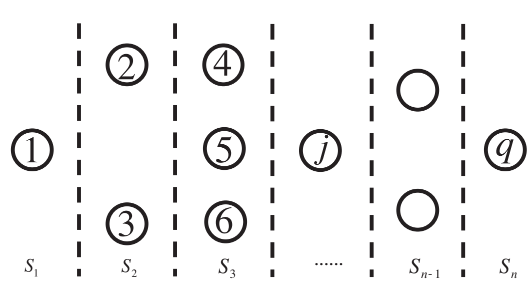

Fig. 2. General model of multistage processes.