The International Journal of Flexible Manufacturing Systems, 15, 167–186, 2003


# ANOVA Method for Variance Component Decomposition and Diagnosis in Batch Manufacturing Processes

JIONGHUA JIN* judy@sie.arizona.edu
HUAIRUI GUO
*Department of Systems & Industrial Engineering, The University of Arizona, Tucson, Arizona 85721-0020, USA*

**Abstract.** In batch manufacturing processes, the total process variation is generally decomposed into batch-by-batch variation and within-batch variation. Since different variation components may be caused by different sources, separation, testing, and estimation of each variance component are essential to the process improvement. Most of the previous SPC research emphasized reducing variations due to assignable causes by implementing control charts for process monitoring. Different from this focus, this article aims to analyze and reduce inherent natural process variations by applying the ANOVA method. The key issue of using the ANOVA method is how to develop appropriate statistical models for all variation components of interest. The article provides a generic framework for decomposition of three typical variation components in batch manufacturing processes. For the purpose of variation root causes diagnosis, the corresponding linear contrasts are defined to represent the possible site variation patterns and the statistical nested effect models are developed accordingly. The article shows that the use of a full factor decomposition model can expedite the determination of the number of nested effect models and the model structure. Finally, an example is given for the variation reduction in the screening conductive gridline printing process for solar battery fabrication.

**Key Words:** analysis of variance, batch manufacturing process, linear contrasts, nested design

## 1. Introduction

Batch manufacturing processes have been widely used in various manufacturing processes, such as wafer fabrication, integrated circuit fabrication, and gridline printing process in solar battery fabrication. In a batch manufacturing process, products are produced batch-by-batch. Thus, total process variations are generally divided into two categories: One is batch-by-batch variation because of the process variability among different batches, and the other is within-batch variation because of the process variability within a batch. Since different variations are usually caused by different root causes, separation and estimation of each variation component is very critical for determining an effective variation reduction strategy in a batch manufacturing process.

Recently, lot of research has been conducted for batch manufacturing process monitoring with the focus on control chart implementation. The emphasized issue is how to develop effective control limits of control charts for monitoring the process condition change, that is, detecting the change of batch-by-batch mean and within-batch variability. The early

*To whom correspondence should be addressed.

approach of the accomplishments is to treat the subgroup statistics within a batch as individual observations. The standard individual control charts are developed on these statistics to monitor batch-by-batch process mean change (Palm, 1992; Porter and Caulcutt, 1992; Wetherill and Brown, 1991; Wheeler and Chambers, 1992). More complicated approaches are to develop control charts to monitor the separated variance components. The initial review of this approach is studied by Woodall and Thomas (1995). Recently, Yashchin (1994) developed CUSUM control charts to monitor the variance components and applied it to the integrated circuit fabrication, in which the variance components representing the lot-to-lot variation and wafer-to-wafer variation, are modeled by the nested random effect models. Roes and Does (1995) developed the monitoring control charts for the silicon wafer manufacturing process, in which a mix-effect model is used to include the fixed effects of the grinding wafer positions in addition to the nested random effect of process runs. In addition, Runger and Fowler (1998) further developed run-to-run control charts with contrasts for semiconductor wafer manufacturing processes. In this approach, the monitored contrasts are formalized based on the engineering knowledge of the site patterns of the potential assignable causes. So, the out-of-control points in the contrast control chart are able to link to the conditions of monitored sites. For multidimensional measurements, multivariate control charts were developed based on PCA (principal component analysis) and PLS (partial least squares) for online monitoring of polymerization reactor process, a stamping process, and a filament extrusion process (Nomikos and MacGregor, 1995; Jin and Shi, 2000, Wurl, Albin, and Shiffer, 2001).

In all above research, the main purpose is to develop effective control charts for detecting process changes. The control limits are determined with the inclusion of the inevitable systematic variability as the inherent natural variability. In fact, prior to the control chart implementation, the first effort is needed in practice to analyze and reduce those initial variation components as much as possible. Then, it is reasonable to justify which systematic variance components have to be included in the control chart development if they eventually cannot be eliminated due to the process constraints. For example, in a wafer manufacturing process, the produced product quality is generally varied over different sites of a silicon wafer. If the sites relating to the nonconforming chips are relatively consistent among wafers over different batches, the first effort should devote to elimination of such a variation component as much as possible. Then, it can determine whether it is still significant and needed to be included as an inherent system variation in the implementation of monitoring control charts. Also, in the example of the screening conductive gridline printing process discussed in Section 3 of the article, the initial setup error of tabletop surface of a Fineline Press machine often leads to an inherent uniformity on the printed gridlines over the printing area. It is recommended to first devote the efforts to reduce or avoid such systematic inherent errors rather than simply accept it as the inherent process variation in the control chart development to avoid getting it worse. So, different from the previous research focus, this article aims to develop a general variation decomposition and analysis methodology for modeling and estimation of inherent variance components in batch manufacturing processes.

The proposed variation modeling and analysis method in the article is based on the general methodology of analysis of variance (ANOVA) method. The critical issue of using the ANOVA method is how to determine appropriate statistical models to describe all decomposed variation components of interest. In the article, a general variation decomposition framework is presented based on the statistical nested effect models for testing and estimation of typical variance components in batch manufacturing processes. The implementation and the effectiveness of the proposed methodology are illustrated in the screening conductive gridline printing process for solar battery fabrication.

The article is organized as follows: after the introduction in Section 1, a general variation decomposition framework of total process variations is provided in Section 2 for all possible nested relationship of variance components of interest. Section 3 gives a brief description of the screening conductive gridline printing process in solar battery fabrication, which will be used as an example in the methodology development. Section 4 discusses the general modeling and analysis procedures using the ANOVA method. The implementation and the effectiveness of the developed methodology are illustrated in Section 5 with a real case study in the gridline printing process. Finally, the article is concluded in Section 6.

## 2. General decomposition framework of total process variations

As is shown in the literature (Woodall and Thomas, 1995; Yashchin, 1994; Roes and Does, 1995; Runger and Fowler, 1998), the process variations in batch manufacturing processes can be generally classified into three types of variations due to the change of three factors or variables, that is, (a) *Factor batch* leading to batch-by-batch variation due to the difference of batches; (b) *Factor sample* inducing sample-by-sample variation representing the difference among samples; and (c) *Factor site* representing the nonuniformity of site-by-site variations. In practice, depending on the potential root causes of each variation component in a particular application, the change of one factor (such as Factor $A$) is usually occurred within another factor (such as Factor $B$). In the analysis of variance (ANOVA), Factor $A$ is nested by Factor $B$. For example, in a wafer fabrication process, the wafer-by-wafer (i.e., sample-by-sample) variation is usually analyzed within a given batch. Thus, factor sample is nested by factor batch. Similarly, in a screening conductive gridline printing process discussed in Section 3, the variance component of factor site is studied within factor run. In this case, factor site is nested by factor batch. Therefore, different nested models should be carefully constructed to appropriately represent those corresponding variance components.

In the article, a general decomposition framework of total process variations is provided in Figure 1, where Layer 1 shows the process variation decomposed by each of three factors, and Layer 2 provides a general list of how each factor's variation is nested by other two factors. Furthermore, to expedite the root cause diagnosis of inherent site variation, the site-by-site variation can be further decomposed into different contrasts defined by linear combinations of site measurements, which is shown in the bottom of Layer 1 in Figure 1. The linear relationships among sites in a contrast should be well defined based on the particular process knowledge such that each contrast corresponds to some meaningful variation patterns of known root causes. When this contrast

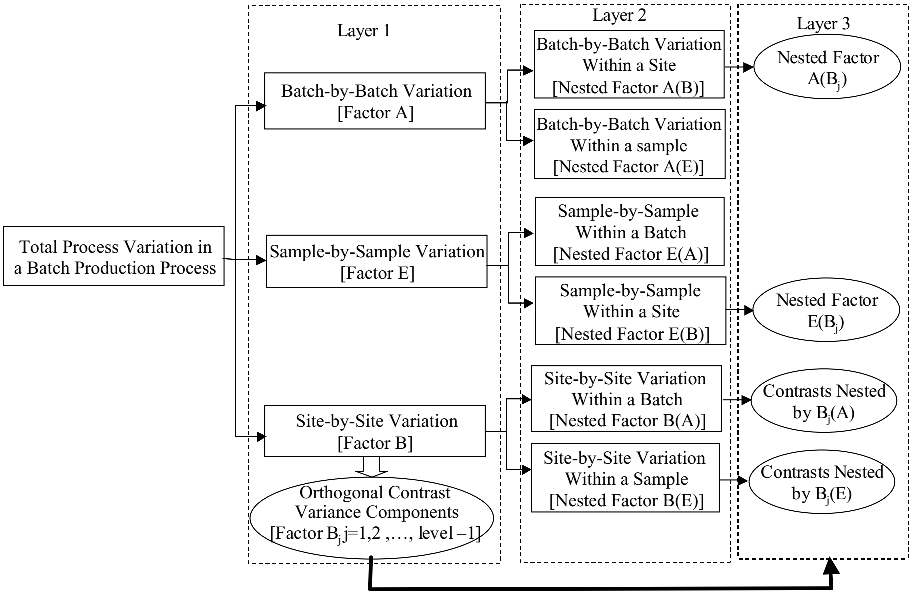


*Figure 1.* General decomposition of variance components.

decomposition of the total site-by-site variation is performed, the variance components, as defined in Layer 2, should be further decomposed under each site contrast rather than the total site variation as shown in Layer 3 of Figure 1. The detailed discussion of this decomposition will be discussed in Section 4 in the screening conductive gridline printing process.

It should be noted that for a particular application, not all listed nest factors necessarily have a clear physical interpretation. The determination of the nest components from this general list is essentially needed for a given application, which usually relies on the process engineering knowledge of the interested variance components and the existence of the possible root causes.

In general, only one statistical model may not be sufficient to describe all interested variation components in Figure 1. So, there is a need to justify how many models are required and what statistical models are adequate for modeling of the selected subset of variance components. In the previous research literatures (Woodall and Thomas, 1995; Yashchin, 1994; Roes and Does, 1995; Runger and Fowler, 1998), there is no discussion on how to use a set of statistical models to fit various decompositions of variance component, because a single statistical model was sufficient to describe the selected monitoring variance components in their applications. In general, the selection of the interested factors (or nested factors) and the determination of the needed statistical models are related to the variation characteristics of a particular application process. As an example, a screening conductive gridline printing process will be used in the article to illustrate the details of the methodology development and implementation procedures.

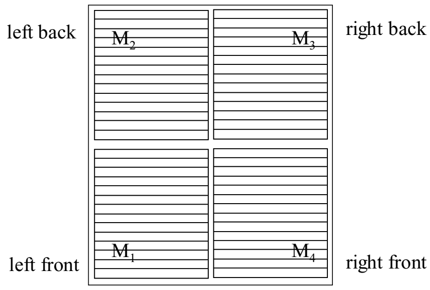


*Figure 2.* Measurement locations on a printed panel.

## 3. Overview of screening conductive gridline printing process

The screening conductive gridline printing process is used to print conductive gridlines on the solar panel to form an electrical circuit, which is a very critical operation in solar battery fabrication. The printing operation is performed panel by panel and each panel is called a *sample*. Generally, the printing tools of a rubber screen and squeegee used in a Fineline Press machine need to be taken out for cleaning after finishing each production run, and then set up again for the next production run. Thus, such a production run cycle under the same tool setup is considered as one-*batch* production.

In the screening conductive gridline printing process, the resistance of printed gridlines is a major concern of the product quality because it significantly affects the solar panel efficiency. It is known that the gridline resistance is affected by the width and height of the printed gridlines. So the spatial uniformity of the width and the height of the printed gridlines is a critical issue for reducing solar panel efficiency variation due to the printing operation. For inspection of the uniformity, four resistance measurements ($M_1$, $M_2$, $M_3$, $M_4$) are taken at four corners of each panel as shown in Figure 2. The variation among these four measurement positions called *sites* is used to represent the spatial uniformity of site-by-site variation.

### *3.1. Description of process variables*

At the beginning of every panel printing, each panel is put on the tabletop of the Fineline Press machine, and the air suction pores distributed on the machine tabletop are used to hold the panel during printing operation. The squeegee moves from front to back to press conductive ink through the rubber screen to print gridlines on the panel. Then, the flood bar moves from back to front to spread ink over the rubber screen ready for printing next panel. The important process variables in the printing process are classified as: (1) tool conditions (screen's tension and missing lines, squeegee's edge shape and hardness), (2) tool setup position (the gap and relative orientation angle between the squeegee and the rubber screen, the alignment of the rubber screen and the squeegee relative to the machine tabletop), and (3) conductive ink material (ink viscosity and composition). Table 1 provides a summary of these process variables and whether they have an effect on each of process variation

*Table 1.* Effect of process variables on potential variations.

| | Batch-by-batch variation | Site-by-site variation | Sample-by-sample variation for a given site within a run |
|---|---|---|---|
| Tool condition | Yes | Yes | Ignore tool degradation within a run |
| Tool setup | Yes | Yes | Assume tool position fixed within a run |
| Conductive ink material | Yes | Ignore ink variation within a run | Assume no new ink added within a run |

components. Since the tool condition degradation is generally very slow, it is reasonable to assume the tool condition is not changed within one production run.

### *3.2. Selection of the interested variation components*

From Table 1, it can be seen that the batch-by-batch variation and site-by-site variation are of the major interest for process variation reduction because they are affected by many assignable process variables or working conditions. The sample-by-sample variation is considered as the inherent process variation especially sampling at a given site and within a production run. Thus, from the general list of variation decomposition in Figure 1, factor batch (Factor $A$) and factor site (Factor $B$) at Layer 1 are considered as the main factors contributing to the process variations. At Layer 2, the possible nested factors can be either $A(B)$ or $B(A)$. Based on the possible existence of potential root causes, three most interested variation components are selected which are contributed by Factor $A$ as well as the nested Factors $A(B)$ and $B(A)$. The details of each variation component and the associated root causes are discussed as follows:

1. Batch-by-batch variation due to the variability of the average of four sites over different runs (Variation $Q_1$ due to Factor $A$): It is mainly represented by the process variability over different batches, such as ink materials, different setup positions of the squeegee orientation angle and the distance of the rubber screen from the printing machine tabletop, and the tool condition change due to squeegee wear and the tension loosing of the rubber screen surface. The process improvement strategy should enhance the inspection of ink property, tool setup, and tool condition at the beginning of each run.
2. Batch-by-batch variation at any of four sites over different runs (Variation $Q_2$ due to the nested factor $A(B)$): It reflects the repeatability of tool setup position at each site over different runs. So, $Q_2$ is caused by different tool setup errors and/or nonuniform tool degradation at each site over different runs. Thus, the reduction of this variation should reduce operator-induced variability and tool setup variability over different runs.
3. Site-by-site variation within a production run (Variation $Q_3$ due to the nested factor $B(A)$): For each production run, the site-by-site variation is mainly caused by the alignment accuracy of the squeegee position relative to the screen positions, and the screen orientation relative to the printing machine tabletop. The strategy for reducing such a

within-run site-by-site variation is to improve the tool setup inspection accuracy at each site and check tool performance uniformity over the printing area.

For above listed three variation components of interest, the ANOVA method will be applied in the following sections to model, test, and estimate these variation components. Then, the effective process improvement efforts can be made accordingly to remove or reduce the corresponding significant variation root causes.

## 4. Variance component analysis using ANOVA

In the article, $y_{ijk}$ is used to denote a measurement taken at site $j(j = 1, \ldots, b)$ of panel sample $k(k = 1, \ldots, n)$ within batch $i$ $(i = 1, \ldots, a)$. In the example of the screening conductive gridline printing process, four measurement sites are fixed for each panel sample, that is, $b = 4$; three panel samples are taken from each batch at each site, that is, $n = 3$; and six batches of production data (i.e. $a = 6$) are selected for initial evaluation of the process variation. The corresponding sum of squares of these three variations ($Q_1$, $Q_2$, and $Q_3$) are represented by

$$SS(Q_1) = SS_A = bn \sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2 \tag{1}$$

$$SS(Q_2) = SS_{A(B)} = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij.} - \bar{y}_{.j.})^2 \tag{2}$$

$$SS(Q_3) = SS_{B(A)} = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij.} - \bar{y}_{i..})^2, \tag{3}$$

Where $y_{i..} = \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}$, $y_{...} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y_{ijk}$, $y_{ij.} = \sum_{k=1}^{n} y_{ijk}$, $y_{.j.} = \sum_{i=1}^{a} \sum_{k=1}^{n} y_{ijk}$, $\bar{y}_{i..} = y_{i..}/nb$, $\bar{y}_{...} = y_{...}/abn$, $\bar{y}_{ij.} = y_{ij.}/n$, $\bar{y}_{.j.} = y_{.j.}/an$. Factor $A$ is a random factor with $a$ levels and Factor $B$ is a fixed factor with $b$ levels.

The critical issue of using the ANOVA method is to develop appropriate linear statistical models for the decomposed variance components of interest. For analyzing the interested sum of squares of $SS_{A(B)}$ and $SS_{B(A)}$ induced by those two nested factors $A(B)$ and $B(A)$, two different nested effect decomposition models are needed and discussed in the following subsections. Section 4.1 will present those two statistical nested models. Section 4.2 gives the statistical testing and estimation of each variance components. To perform root cause determination of the site variation, a further decomposition of the site variation by three linear contrasts is presented in details in Section 4.3.

### *4.1. Statistical modeling of the interested variation components*

A linear two-stage nested design model can be used to represent the interested variance components $Q_1$ and $Q_3$ induced by Factor batch ($A$) and the nested Factor site ($B$) within the batch as follows:

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \varepsilon_{(ij)k} \begin{cases} i = 1, \ldots, a \\ j = 1, \ldots, b \ , \\ k = 1, \ldots, n \end{cases} \tag{4}$$

where $\mu$ is the overall mean of all factors at all levels, $\tau_i$ is the $i$th level effect of Factor $A$, $\beta_{j(i)}$ is the $j$th level effect of Factor $B$ nested under the $i$th level of Factor $A$, and $\varepsilon_{ijk}$ is a random model error with $\varepsilon_{ijk} \sim NID(0, \sigma^2)$. The variance $\sigma^2$ is assumed to be constant and independent of all levels of other factors. Since Factor $A$ is a random factor, it is assumed $\tau_i \sim NID(0, \sigma_\tau^2)$. Although Factor $B$ is a fixed factor, the nested factor of $B$-within-$A$ is a random factor. So, it is assumed that $\beta_{j(i)} \sim NID(0, \sigma^2_{\beta(\tau)})$. From model (4), the total sum of squares can be decomposed as follows called *Decomposition I* in the article:

$$SS_T = SS(Q_1) + SS(Q_3) + SS_E = SS_A + SS_{B(A)} + SS_E \tag{5}$$

Similarly, a different linear two-stage nested design model is needed to represent the interested sum of squares of $Q_2$ induced by Factor batch ($A$) nested by Factor site ($B$) as follows:

$$y_{ijk} = \mu + \tau_{i(j)} + \beta_j + \varepsilon_{(ij)k} \begin{cases} i = 1, \ldots, a \\ j = 1, \ldots, b \\ k = 1, \ldots, n \end{cases} \tag{6}$$

Now, Factor $A$ is random, the nested factor of $A$-within-$B$ is still random, thus, $\tau_{i(j)} \sim NID(0, \sigma^2_{\tau(\beta)})$. However, Factor $B$ is a fixed factor satisfying the condition of $\sum_{j=1}^{b} \beta_j = 0$. In this case, the total sum of squares is decomposed as follows called *Decomposition II* in the article.

$$SS_T = SS_B + SS(Q_2) + SS_E = SS_B + SS_{A(B)} + SS_E \tag{7}$$

### 4.2. *Statistical testing and estimation of the variance components*

In ANOVA analysis, it needs to construct appropriate test statistics to check the significant level of each factor. The $F$-test statistics are obtained as equations ((8)–(10)) based on the relationship of the expected mean squares of the variance components (Montegomery, 1997),

$$\frac{MS_A}{MS_E} \sim F(df_A, df_E) \tag{8}$$

$$\frac{MS_{A(B)}}{MS_E} \sim F\big(df_{A(B),} df_E\big) \tag{9}$$

$$\frac{MS_{B(A)}}{MS_E} \sim F\big(df_{B(A)}, df_E\big), \tag{10}$$

where $MS_A$, $MS_{A(B)}$, and $MS_{B(A)}$ are the mean squares of the corresponding factors described by each subscript. $df_A$, $df_{A(B)}$, and $df_{B(A)}$ are the corresponding degrees of the freedom. The detailed analysis of each item is shown in the expected mean squares tables (Tables 8 and 10) and the analysis of variance tables (Tables 9 and 10) in the Appendix. The estimates of these variance components are also obtained as

$$\hat{\sigma}^2_{Q_1} = \hat{\sigma}^2_\tau = (MS_A - MS_E)/bn \tag{11}$$

$$\hat{\sigma}^2_{Q_2} = \hat{\sigma}^2_{\tau(\beta)} = \left(MS_{A(B)} - MS_E\right)\big/n \tag{12}$$

$$\hat{\sigma}^2_{Q_3} = \hat{\sigma}^2_{\beta(\tau)} = \left(MS_{B(A)} - MS_E\right)\big/n \tag{13}$$

*4.3. Site variability decomposition and diagnosis based on linear contrasts*

Both variation components of $Q_2$ and $Q_3$ are related to the site variability. If such variation components are significant, removal or reducing of the associated root causes are generally desired. From Table 1, it can be seen that the site variability is mainly related to the tool conditions and tool setups. Thus, identification of site variation patterns can expedite the diagnosis of variation root causes for process improvement.

Based on the engineering knowledge of potential fault patterns in the screening conductive gridline printing process, the total spatial variation among four sites can be further decomposed into three typical variation patterns, which are reflected by the differences of the front and back site (between M1&M4 and M2&M3), the left and right site (between M1&M2 and M3&M4), and the diagonals (between M1&M3 and M2&M4) as shown in Figure 2. The statistical test and the estimates of these three variation patterns can provide a useful guideline on how to adjust the tool position to reduce tool setup errors.

For the purpose of analyzing the contribution of each variation pattern, three linear contrasts of $B^{FB}$, $B^{LR}$, and $B^D$ are defined to represent each contrast effect of front-back sites, left-right sites, and the diagonal respectively, that is,

$$\beta^{FB}_{ik} = (y_{i1k} + y_{i4k} - y_{i2k} - y_{i3k})/2 \tag{14}$$

$$\beta^{LR}_{ik} = (y_{i1k} + y_{i2k} - y_{i3k} - y_{i4k})/2 \tag{15}$$

$$\beta^{D}_{ik} = (y_{i1k} + y_{i3k} - y_{i2k} - y_{i4k})/2 \tag{16}$$

The constant 2 in the denominator is used to normalize the effect of each contrast. It is known that two contrasts with linear coefficients $\{c_i\}$ and $\{d_i\}$ are called orthogonal contrasts if the condition of $\sum_{i=1}^{b} c_i d_j = 0$ is satisfied. From equations (14)-(16), it can be seen that the linear coefficients of those three contrasts $\beta^{FB}_{ik}$, $\beta^{LR}_{ik}$, and $\beta^{D}_{ik}$ are [0.5 −0.5 −0.5 0.5], [0.5 0.5 −0.5 −0.5], and [0.5 −0.5 0.5 −0.5], respectively. So, it can be proved that the defined three contrasts are orthogonal to each other. The following analysis will show how to develop appropriate statistical models to meet the need of analyzing these contrast variations for *decomposition I* of equation (5) and *decomposition II* of equation (7), respectively.

***4.3.1. Analysis of contrast variance components for decomposition I***

*4.3.1.1. Variation decomposition I represented by linear contrasts.* It is known that the total sum of squares $SS_T$ can be represented by

$$SS_T = \sum_{i=1}^{a}\sum_{k=1}^{n}[(y_{i1k} - \bar{y}_{...})^2 + (y_{i2k} - \bar{y}_{...})^2 + (y_{i3k} - \bar{y}_{...})^2 + (y_{i4k} - \bar{y}_{...})^2] \tag{17}$$

It can be shown that

$$y_{i1k} - \bar{y}_{...} = (\bar{y}_{i..} - \bar{y}_{...}) + \left(\frac{\bar{\beta}_{i.}^{FB} + \bar{\beta}_{i.}^{LR} + \bar{\beta}_{i.}^{D}}{2}\right) + (y_{i1k} - \bar{y}_{i1.}) \tag{18}$$

So, the sum of squares of equation (18) is:

$$\sum_{i=1}^{a}\sum_{k=1}^{n}(y_{i1k} - \bar{y}_{...})^2 = n\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2 + \frac{n}{4}\sum_{i=1}^{a}\left[\left(\bar{\beta}_{i.}^{FB}\right)^2 + \left(\bar{\beta}_{i.}^{LR}\right)^2 + \left(\bar{\beta}_{i.}^{D}\right)^2\right] + \sum_{i=1}^{a}\sum_{k=1}^{n}(y_{i1k} - \bar{y}_{i1.})^2 \tag{19}$$

The derivation from equation (18) to equation (19) utilizes the fact that all four items of $\bar{y}_{i..}$, $\bar{\beta}_{i.}^{FB}$, $\bar{\beta}_{i.}^{LR}$, and $\bar{\beta}_{i.}^{D}$ are orthogonal to each other.

Similarly, for other three variation components of equation (17), it can be obtained that

$$y_{i2k} - \bar{y}_{...} = (\bar{y}_{i..} - \bar{y}_{...}) + \frac{-\bar{\beta}_{i.}^{FB} + \bar{\beta}_{i.}^{LR} - \bar{\beta}_{i.}^{D}}{2} + y_{i2k} - \bar{y}_{i2.} \tag{20}$$

$$y_{i3k} - \bar{y}_{...} = (\bar{y}_{i..} - \bar{y}_{...}) + \frac{-\bar{\beta}_{i.}^{FB} + \bar{\beta}_{i.}^{LR} + \bar{\beta}_{i.}^{D}}{2} + y_{i3k} - \bar{y}_{i3.} \tag{21}$$

$$y_{i4k} - \bar{y}_{...} = (\bar{y}_{i..} - \bar{y}_{...}) + \frac{\bar{\beta}_{i.}^{FB} - \bar{\beta}_{i.}^{LR} - \bar{\beta}_{i.}^{D}}{2} + y_{i4k} - \bar{y}_{i4.} \tag{22}$$

Thus, the total sum of squares of equation (17) can be represented in terms of the nested contrasts as

$$SS_T = 4n\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2 + n\sum_{i=1}^{a}\left[\left(\bar{\beta}_{i.}^{FB}\right)^2 + \left(\bar{\beta}_{i.}^{LR}\right)^2 + \left(\bar{\beta}_{i.}^{D}\right)^2\right] + \sum_{i=1}^{a}\sum_{j=1}^{4}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2 = SS_A + SS_{B^{FB}(A)} + SS_{B^{LR}(A)} + SS_{B^{D}(A)} + SS_E \tag{23}$$

It concludes that the original total site variation nested by Factor run are decomposed into three orthogonal contrast variations nested by Factor run, that is, $SS_{B(A)} = SS_{B^{FB}(A)} + SS_{B^{LR}(A)} + SS_{B^{D}(A)}$. This result shows that for a factor with level $b$, its total variation nested by another factor can still be decomposed into the variations of $b - 1$ orthogonal contrasts under the same nest factor.

*4.3.1.2. Representation of contrast effects by factor effects.* In fact, three contrasts defined above can also be further represented by two factors $V$ and W with two fixed levels, which are defined as $V = 1$ representing the front site, $V = -1$ representing the back site; $W = 1$ representing the left site, and $W = -1$ representing the right site, that is,

$$\nu_{i1k} = (y_{i1k} + y_{i4k})/\sqrt{2}, \quad \nu_{i2k} = (y_{i2k} + y_{i3k})/\sqrt{2} \tag{24}$$

$$\omega_{i1k} = (y_{i1k} + y_{i2k})/\sqrt{2}, \quad \omega_{i2k} = (y_{i3k} + y_{i4k})/\sqrt{2} \tag{25}$$

$\sqrt{2}$ is used to normalize each factor effect. The equivalence of the defined new factors to the contrasts can be seen by the following proof through comparing the results in equation (29) with that in equation (23).

$$\begin{aligned} SS_{V(A)} &= n\sum_{i=1}^{a}[(\bar{\nu}_{i1.} - \bar{\nu}_{i..})^2 + (\bar{\nu}_{i2.} - \bar{\nu}_{i..})^2] \\ &= n\sum_{i=1}^{a}\left[\left(\frac{\bar{y}_{i1.} + \bar{y}_{i4.}}{\sqrt{2}} - \frac{\bar{y}_{i1.} + \bar{y}_{i2.} + \bar{y}_{i3.} + \bar{y}_{i4.}}{2\sqrt{2}}\right)^2\right. \\ &\quad \left. + \left(\frac{\bar{y}_{i2.} + \bar{y}_{i3.}}{\sqrt{2}} - \frac{\bar{y}_{i1.} + \bar{y}_{i2.} + \bar{y}_{i3.} + \bar{y}_{i4.}}{2\sqrt{2}}\right)^2\right] \\ &= n\sum_{i=1}^{a}\left(\frac{\bar{y}_{i1.} + \bar{y}_{i4.} - \bar{y}_{i2.} - \bar{y}_{i3.}}{2}\right)^2 = SS_{B^{FB}(A)} \end{aligned} \tag{26}$$

Similarly, the sum of squares of the nested factor W as well as the nested interaction of Factor $V$ and Factor $W$ are

$$\begin{aligned} SS_{W(A)} &= n\sum_{i=1}^{a}\left[\left(\frac{\bar{y}_{i1.} + \bar{y}_{i2.}}{\sqrt{2}} - \frac{\bar{y}_{i1.} + \bar{y}_{i2.} + \bar{y}_{i3.} + \bar{y}_{i4.}}{2\sqrt{2}}\right)^2\right. \\ &\quad \left. + \left(\frac{\bar{y}_{i3.} + \bar{y}_{i4.}}{\sqrt{2}} - \frac{\bar{y}_{i1.} + \bar{y}_{i2.} + \bar{y}_{i3.} + \bar{y}_{i4.}}{2\sqrt{2}}\right)^2\right] \end{aligned} \tag{27}$$

$$\begin{aligned} SS_{V\times W(A)} &= n\sum_{i=1}^{a}\left[\left(\frac{\bar{y}_{i1.} + \bar{y}_{i3.}}{\sqrt{2}} - \frac{\bar{y}_{i1.} + \bar{y}_{i2.} + \bar{y}_{i3.} + \bar{y}_{i4.}}{2\sqrt{2}}\right)^2\right. \\ &\quad \left. + \left(\frac{\bar{y}_{i2.} + \bar{y}_{i4.}}{\sqrt{2}} - \frac{\bar{y}_{i1.} + \bar{y}_{i2.} + \bar{y}_{i3.} + \bar{y}_{i4.}}{2\sqrt{2}}\right)^2\right] \end{aligned} \tag{28}$$

It shows that $SS_{W(A)} = SS_{LR(A)}$ and $SS_{[V\times W](A)} = SS_{D(A)}$. Thus, equation (23) can be represented as

$$SS_T = SS_A + SS_{V(A)} + SS_{W(A)} + SS_{[V\times W](A)} + SS_E \tag{29}$$

*Remark.* An important conclusion is that the total process variation is contributed by three factors as random Factor $A$ with the level equal to the number of runs and two fixed factors of $V$ and $W$ with the level equal to two. The main effects of Factors $V$ and $W$ as well as the effect of their interaction $V \times W$ are equivalent to the contrast effects of $B^{FB}$, $B^{LR}$, and $B^{D}$, respectively. This conclusion can be directly used to expedite the development of statistical model for the contrast variations under decomposition II, which is discussed in Section 4.3.2.

*4.3.1.3. Statistical testing and estimation of the decomposed site variance components.* Statistical modeling of the decomposition of equation (29) can be represented by an equivalent model of the two-stage nested design of three factors as

$$y_{ijk} = \mu + \tau_i + \nu_{p(i)} + \omega_{q(i)} + (\nu\omega)_{pq(i)} + \varepsilon_{(ipq)k} \begin{cases} i = 1, \ldots, a \\ p, q = 1, 2 \\ k = 1, \ldots, n \end{cases} \quad (30)$$

In the model, $\tau_i$ is the effect of run $i$. $\nu_{p(i)}$ is the effect of Factor $V$ at level $p$ nested by Factor $A$ at level $i$, that is, the effect of the front-back contrast within run $i$. Similarly, $\omega_{q(i)}$ is the effect of the left-right contrast within run $i$, $(\nu\omega)_{pq(i)}$ is the effect of the diagonal contrast within run $i$. So, the main effect of $\tau_i$ is the same as the original model of equation (4), the nested effects of $\nu_{p(i)}$, $\omega_{q(i)}$, and $(\nu\omega)_{pq(i)}$ represent the within-run variations induced by different site variation patterns. Now, using the decomposition model of equation (30), the $F$-tests of the nested contrast factors are developed in equations (31)–(33) based on the expected mean squares table (Table 12) and the analysis of variance table (Table 13) in the Appendix.

$$\frac{MS_{V(A)}}{MS_E} \sim F(a, df_E) \quad (31)$$

$$\frac{MS_{W(A)}}{MS_E} \sim F(a, df_E) \quad (32)$$

$$\frac{MS_{[V\times W](A)}}{MS_E} \sim F(a, df_E) \quad (33)$$

The corresponding nested variance components are estimated by

$$\hat{\sigma}^2_{Q_3^V} = \hat{\sigma}^2_{\nu(\tau)} = \left(MS_{V(A)} - MS_E\right)/2n \quad (34)$$

$$\hat{\sigma}^2_{Q_3^W} = \hat{\sigma}^2_{\omega(\tau)} = \left(MS_{W(A)} - MS_E\right)/2n \quad (35)$$

$$\hat{\sigma}^2_{Q_3^{V\times W}} = \hat{\sigma}^2_{\nu\times\omega(\tau)} = \left(MS_{[V\times W](A)} - MS_E\right)/n \quad (36)$$

***4.3.2. Analysis of contrast variations for decomposition II.*** For decomposition II, the original nested Factor $A(B)$ is further decomposed into three items, that is, Factor $A$ nested by three contrasts can be equivalent to Factor $A$ nested by Factor $V$, Factor $W$, and their interaction $V \times W$ as $A(V)$, $A(W)$, and $A(V \times W)$.

From the conclusion in the remark of Section 4.3.1, it is known that the total process variation can be generally considered as the effect of these three factors ($A$, $V$, and $W$) and their interactions. Thus, a full factor decomposition of the total sum of squares of three factors can be generally expressed as

$$SS_T = SS_A + SS_V + SS_W + SS_{A\times V} + SS_{A\times W} + SS_{V\times W} + SS_{A\times V\times W} + SS_E \quad (37)$$

From the nested design, it is known that $SS_{A(V)} = SS_A + SS_{A\times V}$, $SS_{A(W)} = SS_A + SS_{A\times W}$, and $SS_{A(V\times W)} = SS_A + SS_{A\times V} + SS_{A\times W} + SS_{A\times V\times W}$. Thus, three decomposition models are needed to represent all these nested factors of $A(V)$, $A(W)$, and $A(V \times W)$, that is,

$$SS_T = SS_{A(V)} + SS_V + SS_W + SS_{V\times W} + SS_{A\times W(V)} + SS_E \quad (38)$$

$$SS_T = SS_{A(W)} + SS_V + SS_W + SS_{V\times W} + SS_{A\times V(W)} + SS_E \quad (39)$$

$$SS_T = SS_{A(V\times W)} + SS_V + SS_W + SS_{V\times W} + SS_E \quad (40)$$

Here, both equations (38) and (39) are corresponding to the standard two-stage nested design models. In fact, it is known that $SS_B = SS_V + SS_W + SS_{V\times W}$. Thus, by comparing equation (40) with equation (7), it can be obtained that $SS_{A(V\times W)} = SS_{A(B)}$. So, there is no need to analyze equation (40) anymore. The statistical models corresponding to equations (38) and (39) are

$$y_{ijk} = \mu + \tau_{i(p)} + \nu_p + \omega_q + (\nu\omega)_{pq} + (\tau\omega)_{iq(p)} + \varepsilon_{(ipq)k} \quad (41)$$

$$y_{ijk} = \mu + \tau_{i(q)} + \nu_p + \omega_q + (\nu\omega)_{pq} + (\tau\omega)_{ip(q)} + \varepsilon_{(ipq)k} \quad (42)$$

From equation (41), the $F$-test for the nested Factor $A(V)$ is obtained based on the expected mean squares table and the analysis of variance table shown in Tables 14 and 15 in the Appendix.

$$\frac{MS_{A(V)}}{MS_E} \sim F(2(a-1), df_E) \quad (43)$$

The corresponding estimation of this variance component is

$$\hat{\sigma}^2_{\mathrm{Q_2}(V)} = \hat{\sigma}^2_{\tau(\nu)} = \big(MS_{A(V)} - MS_E\big)\big/2n \quad (44)$$

A similar $F$-test can be conducted for the nested Factor $A(W)$ and estimation of $\hat{\sigma}^2_{\mathrm{Q_2}(W)}$.

## 5. Case study and results

A case study was conducted by initially analyzing six runs of old production data to provide some suggestions for process improvement. A follow-up validation is further made through collecting and analyzing another six runs of new production data after the process takes the corresponding improvement.

First, the variance component analysis is conducted based on two factors, that is, a random run factor $A$ and a fixed site factor $B$. If the effect of Factor $B$ is significant, a further diagnostic analysis is made for the contrast variation components through three factors, that is, run factor $A$, front-back factor $V$, and left-right factor $W$. For the old production data, Table 2 shows the sum of squares of the full decomposition model based on Factor $A$ and Factor $B$ with the levels of $a = 6$ and $b = 4$, respectively.

*Table 2.* Sum of squares of the full factor decomposition for two factors.

| | $A$ | $B$ | $A \times B$ | $E$ | Total |
|---|---|---|---|---|---|
| $SS$ | 0.045664 | 0.017633 | 0.049859 | 0.085920 | 0.199075 |
| $df$ | 5 | 3 | 15 | 48 | 71 |

*Table 3.* $F$-test of variance components for old process.

| Variation source | Sum of squares ($SS$) | Degree of freedom ($df$) | Mean square ($MS$) | $F$-test ($F_0$) | $P$ |
|---|---|---|---|---|---|
| $A \Rightarrow Q_1$ | $SS_A = 0.045664$ | $df_A = a - 1$ $= 5$ | $MS_A = \frac{SS_A}{df_A}$ $= 0.009133$ | $\frac{MS_A}{MS_E} = 5.10$ | 0.00079 |
| $A(B) \Rightarrow Q_2$ | $SS_{A(B)} = SS_A + SS_{A\times B}$ $= 0.095523$ | $df_{A(B)} = b(a-1)$ $= 20$ | $MS_{A(B)} = \frac{SS_{A(B)}}{df_{A(B)}}$ $= 0.004776$ | $\frac{MS_{A(B)}}{MS_E} = 2.67$ | 0.0028 |
| $B(A) \Rightarrow Q_3$ | $SS_{B(A)} = SS_B + SS_{A\times B}$ $= 0.067492$ | $df_{B(A)} = a(b-1)$ $= 18$ | $MS_{B(A)} = \frac{SS_{B(A)}}{df_{B(A)}}$ $= 0.003750$ | $\frac{MS_{B(A)}}{MS_E} = 2.09$ | 0.021 |
| Error $E$ | $SS_E = 0.08592$ | $df_E = ab(n-1)$ $= 48$ | $MS_E = \frac{SS_E}{df_E}$ $= 0.00179$ | | |

The analysis of the interested variance components $Q_1$, $Q_2$, and $Q_3$ is made and summarized in Table 3, where $P$-value indicates the significant level of $F$-tests.

From this analysis, it can be seen that the batch mean variation $Q_1$ is significant with a Type I error not larger than 0.00079. Also, the site effect is significant, which is indicated by the $F$-test for $Q_2$ with a Type I error not larger than 0.0028, and for $Q_3$ with a Type I error not larger than 0.021. Therefore, the process improvement should focus on both batch mean variation reduction and site variation reduction. The estimation of each variance components is calculated from equations (11)–(13) as, $\hat{\sigma}^2_{Q_1} = 0.00061$, $\hat{\sigma}^2_{Q_2} = 0.001$, and $\hat{\sigma}^2_{Q_3} = 0.00065$.

To effectively find the root causes of site variations, the analysis of contrast variance components is further conducted. The sum of squares of the full decomposition model in equation (37) is summarized in Table 4. The further statistical $F$-tests of the contrast variation components by using the nested models of equations (38) and (39) are summarized in Table 5.

It is clear that the run-by-run variations nested by Factor $V$ and Factor $W$ [$A(V)$ and $A(W)$] are all significant. Therefore, the tool setup repeatability should be enhanced to reduce the run-by-run variation. However, for the within-run variation, only the effect of the back-front contrast $V(A)$ is significant with a Type I error not more than 0.0016. Thus,

*Table 4.* Sum of squares based on the full decomposition of three factors.

| | $A$ | $V$ | $W$ | $V \times W$ | $A \times V$ | $A \times W$ | $A \times V \times W$ | $E$ | Total |
|---|---|---|---|---|---|---|---|---|---|
| $SS$ | 0.045664 | 0.007653 | 0.003906 | 0.006074 | 0.038094 | 0.009663 | 0.002101 | 0.085920 | 0.199075 |
| $df$ | 5 | 1 | 1 | 1 | 5 | 5 | 5 | 48 | 71 |

*Table 5.* $F$-tests for the contrast variations.

| Source of variation | Sum of squares | Mean square | $F$-test | $P$ |
|---|---|---|---|---|
| $Q_2(B^{FB}) \Rightarrow A(B^{FB}) \Rightarrow A(V)$ | $SS_{A(V)} = SS_A + SS_{A\times V} = 0.083758$ | $MS_{A(V)} = \frac{SS_{A(V)}}{2(a-1)} = 0.008376$ | $\frac{MS_{A(V)}}{MS_E} = 4.68$ | 0.00016 |
| $Q_2(B^{LR}) \Rightarrow A(B^{LR}) \Rightarrow A(W)$ | $SS_{A(W)} = SS_A + SS_{A\times W} = 0.055327$ | $MS_{A(W)} = \frac{SS_{A(W)}}{2(a-1)} = 0.0055327$ | $\frac{MS_{A(W)}}{MS_E} = 3.09$ | 0.0041 |
| $Q_3^{FB}(A) \Rightarrow B^{FB}(A) \Rightarrow V(A)$ | $SS_{V(A)} = SS_V + SS_{V\times A} = 0.045747$ | $MS_{V(A)} = \frac{SS_{V(A)}}{a} = 0.007625$ | $\frac{MS_{V(A)}}{MS_E} = 4.26$ | 0.0016 |
| $Q_3^{LR}(A) \Rightarrow B^{LR}(A) \Rightarrow W(A)$ | $SS_{W(A)} = SS_W + SS_{W\times A} = 0.013569$ | $MS_{W(A)} = \frac{SS_{W(A)}}{a} = 0.002262$ | $\frac{MS_{W(A)}}{MS_E} = 1.26$ | 0.29 |
| $Q_3^{D}(A) \Rightarrow B^{D}(A) \Rightarrow [W \times V](A)$ | $SS_{[V\times W](A)} = SS_{V\times W} + SS_{A\times V\times W} = 0.008175$ | $MS_{V\times W(A)} = \frac{SS_{[V\times W](A)}}{a} = 0.001362$ | $\frac{MS_{[V\times W](A)}}{MS_E} = 0.76$ | 0.60 |
| Error $E$ | $SS_E = 0.08592$ | $MS_E = \frac{SS_E}{4a(n-1)} = 0.001790$ | | |

*Table 6.* Sum of squares based on the full decomposition model of the new process.

| | $A$ | $B$ | $A \times B$ | $E$ | Total |
|---|---|---|---|---|---|
| $SS$ | 0.008621 | 0.002841 | 0.021428 | 0.08425 | 0.11714 |
| $df$ | 5 | 3 | 15 | 48 | 71 |

the process improvement strategy for reducing the within-run variation should mainly focus on reducing the back-front error in every run. It turns out that the fixed machine tabletop position error is the major contribution of this within-run site variation. The estimates of each significant variance component can be calculated from equations (34) and (44) yielding $\hat{\sigma}^2_{\nu(\tau)} = 0.00097$, $\hat{\sigma}^2_{\tau(\nu)} = 0.0011$, and $\hat{\sigma}^2_{\tau(\omega)} = 0.00062$.

After a careful adjustment of the printing machine tabletop and improving the inspection of tool setup alignment and ink variability, six runs of new process data are collected again for validation analysis. The variance component analysis based on the full factor

*Table 7.* $F$-tests of variance components for the new process.

| Variation source | Sum of squares ($SS$) | Degree of freedom ($df$) | Mean square ($MS$) | $F$-test ($F_0$) | $P$ |
|---|---|---|---|---|---|
| $A \Rightarrow Q_1$ | $SS_A = 0.008621$ | $df_A = a - 1 = 5$ | $MS_A = 0.001724$ | $\frac{MS_A}{MS_E} = 0.98$ | 0.44 |
| $A(B) \Rightarrow Q_2$ | $SS_{A(B)} = SS_A + SS_{A\times B} = 0.030049$ | $df_{A(B)} = b(a-1) = 20$ | $MS_{A(B)} = 0.001502$ | $\frac{MS_{A(B)}}{MS_E} = 0.86$ | 0.64 |
| $B(A) \Rightarrow Q_3$ | $SS_{B(A)} = SS_B + SS_{A\times B} = 0.024268$ | $df_{B(A)} = a(b-1) = 18$ | $MS_{B(A)} = 0.001348$ | $\frac{MS_{B(A)}}{MS_E} = 0.77$ | 0.72 |
| Error $E$ | 0.08425 | $df_E = ab(n-1) = 48$ | $MS_E = 0.00175$ | | |

decomposition model of this new process is shown in Table 6. The $F$-test results of the nested variance components are given in Table 7. Now, it shows there is no significant factor anymore. The process improvement efficiency can be indicated by the percentage of the total variation reduction as:

$$\eta = \frac{Var(old) - Var(new)}{Var(old)} \times 100\% = \frac{SS_T(old) - SS_T(new)}{SS_T(old)} \times 100\% = 41.13\%$$

## 6. Conclusion

In the article, the variation decomposition and analysis for batch manufacturing processes has been conducted by using the ANOVA method. A generic framework is provided for decomposition of three typical variation components in batch manufacturing processes. For a significant site variability, three diagnostic contrast components are defined for root cause identification of the potential site variation patterns. When the interested variations are induced by more than three factors, it generally needs many different nested models for various variation component analyses. The use of a full factor decomposition model to determine the needed nested models is suggested in the article, which can expedite the statistical model development. Finally, the article illustrates the effectiveness of the proposed method in the screening conductive gridline printing process.

Although the proposed variation analysis method is illustrated in a specific application of the printing process, the proposed analysis method can be applied to other batch manufacturing processes where separation of the batch-by-batch variation and the within-batch variation is needed. Also, when the potential variation patterns can be well defined by linear contrasts, the usage of the diagnostic contrasts can expedite the variation root cause determination. If the contrast effects cannot be transferred into factors' effect, a direct decomposition of the sum of squares of the contrast components is needed to the development of the nested effect models. It should be noted that when the interested variance components include sample-to-sample variation, different models can be developed in the same way for independent samples. However, when the samples are dependent, the proposed method cannot be directly applied. This issue will be considered as the further research in future.

## Appendix

For decomposition I of equation (5), Factor $A$ and Factor $B$-within-$A$ are all random. The determination of the expected mean squares table is shown in Table 8 based on the general

*Table 8.* Expected mean square for decomposition I of equation (5).

| Factor | R<br>$a$<br>$i$ | F<br>$b$<br>$j$ | R<br>$n$<br>$k$ | Expected mean square |
|---|---|---|---|---|
| $\tau_i$ | 1 | $b$ | $n$ | $E(MS_A) = \sigma^2 + bn\sigma_\tau^2$ |
| $\beta_{j(i)}$ | 1 | 0 | $n$ | $E(MS_{B(A)}) = \sigma^2 + n\sigma_{\beta(\tau)}^2$ |
| $\varepsilon_{(ij)k}$ | 1 | 1 | 1 | $E(MS_E) = \sigma^2$ |

*Table 9.* Analysis of variance table for decomposition I of equation (5).

| Source of variation | Sum of squares (*SS*) | Degree of freedom ($df$) | Mean square (*MS*) | $F$-test ($F_0$) |
|---|---|---|---|---|
| Random factor $A \Rightarrow Q_1$ | $SS_A = bn \sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2$ | $df_A = a - 1$ | $MS_A = \frac{SS_A}{df_A}$ | $F_0^A = \frac{MS_A}{MS_E}$ |
| Random Factor $B(A) \Rightarrow Q_3$ | $SS_{B(A)} = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij.} - \bar{y}_{i..})^2$ | $df_{B(A)} = a(b-1)$ | $MS_{B(A)} = \frac{SS_{B(A)}}{df_{B(A)}}$ | $F_0^{B(A)} = \frac{MS_{B(A)}}{MS_E}$ |
| Error $E$ | $SS_E = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{ij.})^2$ | $df_E = ab(n-1)$ | $MS_E = \frac{SS_E}{df_E}$ | |
| Total | $SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{...})^2$ | $df_T = abn - 1$ | | |

*Table 10.* Expected mean square for decomposition II of equation (7).

| Factor | R<br>$a$<br>$i$ | F<br>$b$<br>$j$ | R<br>$n$<br>$k$ | Expected mean square |
|---|---|---|---|---|
| $\tau_{i(j)}$ | 1 | 1 | $n$ | $E(MS_{A(B)}) = \sigma^2 + n\sigma^2_{\tau(\beta)}$ |
| $\beta_j$ | $a$ | 0 | $n$ | $E(MS_B) = \sigma^2 + n\sigma^2_{\tau(\beta)} + (an \sum_{j=1}^{b} \beta_j^2)/(b-1)$ |
| $\varepsilon_{(ij)k}$ | 1 | 1 | 1 | $E(MS_E) = \sigma^2$ |

*Table 11.* Analysis of variance table for decomposition II of equation (7).

| Source of variation | Sum of squares (*SS*) | Degree of freedom ($df$) | Mean square (*MS*) | $F$-test ($F_0$) |
|---|---|---|---|---|
| Random factor $A(B) \Rightarrow Q_2$ | $SS_{A(B)} = n \sum_{i=1}^{a} \sum_{j=1}^{b} \times (\bar{y}_{ij.} - \bar{y}_{.j.})^2$ | $df_{A(B)} = b(a-1)$ | $MS_{A(B)} = \frac{SS_{A(B)}}{df_{A(B)}}$ | $F_0^{A(B)} = \frac{MS_{A(B)}}{MS_E}$ |
| Fixed factor $B$ | $SS_B = an \sum_{j=1}^{b} (\bar{y}_{.j.} - \bar{y}_{...})^2$ | $df_B = b - 1$ | $MS_B = \frac{SS_B}{df_B}$ | $F_0^B = \frac{MS_B}{MS_{A(B)}}$ |
| Error $E$ | $SS_E = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{ij.})^2$ | $df_E = ab(n-1)$ | $MS_E = \frac{SS_E}{df_E}$ | |
| Total | $SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{...})^2$ | $df_T = abn - 1$ | | |

rules discussed in Montegomery (1997). In the table, the first row indicates whether the factor is random (R) or fixed (F), the second row indicates the level of each factor, and the third row is the index of the factor level. A summary of the analysis of variance is shown in Table 9.

Similar analysis is performed for decomposition II of equation (7), where Factor $A$-within-$B$ is a random factor but Factor $B$ is a fixed factor. The determination of the expected mean squares table is shown in Table 10. A corresponding summary of the analysis of variances is given in Table 11.

*Table 12.* Expected mean squares for contrast decomposition I of equation (29).

| Factor | R<br>$a$<br>$i$ | F<br>2<br>$p$ | F<br>2<br>$q$ | R<br>$n$<br>$k$ | Expected mean square |
|---|---|---|---|---|---|
| $\tau_i$ | 1 | 2 | 2 | $n$ | $E(MS_A) = \sigma^2 + 4n\sigma_\tau^2$ |
| $\beta_{(\tau)}^{FB} \Rightarrow \nu_{p(i)}$ | 1 | 0 | 2 | $n$ | $E(MS_{V(A)}) = \sigma^2 + 2n\sigma_{\nu(\tau)}^2$ |
| $\beta_{(\tau)}^{LR} \Rightarrow \omega_{q(i)}$ | 1 | 2 | 0 | $n$ | $E(MS_{W(A)}) = \sigma^2 + 2n\sigma_{\omega(\tau)}^2$ |
| $\beta_{(\tau)}^{D} \Rightarrow (\nu\omega)_{pq(i)}$ | 1 | 0 | 0 | $n$ | $E(MS_{V\times W(A)}) = \sigma^2 + n\sigma_{\nu\times\omega(\tau)}^2$ |
| $\varepsilon_{(ipq)k}$ | 1 | 1 | 1 | 1 | $E(MS_E) = \sigma^2$ |

*Table 13.* Analysis of variance table for contrast decomposition I of equation (29).

| Source of variation | Sum of squares | Mean square | $F$ test |
|---|---|---|---|
| Random factor $A \Rightarrow Q_1$ | $SS_A = bn\sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2$ | $MS_A = \frac{SS_A}{a-1}$ | $F_0^A = \frac{MS_A}{MS_E}$ |
| Random factor $B^{FB}(A) \Rightarrow V(A)$ | $SS_{V(A)} = n\sum_{i=1}^{a} \times (\frac{\bar{y}_{i1.}+\bar{y}_{i4.}-\bar{y}_{i2.}-\bar{y}_{i3.}}{2})^2$ | $MS_{V(A)} = \frac{SS_{V(A)}}{a}$ | $F_0^{V(A)} = \frac{MS_{V(A)}}{MS_E}$ |
| Random factor $B^{LR}(A) \Rightarrow W(A)$ | $SS_{W(A)} = n\sum_{i=1}^{a} \times (\frac{\bar{y}_{i1.}+\bar{y}_{i2.}-\bar{y}_{i3.}-\bar{y}_{i4.}}{2})^2$ | $MS_{W(A)} = \frac{SS_{W(A)}}{a}$ | $F_0^{W(A)} = \frac{MS_{W(A)}}{MS_E}$ |
| Random factor $B^D(A) \Rightarrow V \times W(A)$ | $SS_{[V\times W](A)} = n\sum_{i=1}^{a} \times (\frac{\bar{y}_{i1.}+\bar{y}_{i2.}-\bar{y}_{i3.}-\bar{y}_{i4.}}{2})^2$ | $MS_{[V\times W](A)} = \frac{SS_{[V\times W](A)}}{a}$ | $F_0^{V\times W(A)} = \frac{MS_{[V\times W](A)}}{MS_E}$ |
| Error $E$ | $SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b=4}\sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{ij.})^2$ | $MS_E = \frac{SS_E}{4a(n-1)}$ | |
| Total | $SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b=4}\sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{...})^2$ | | |

*Table 14.* Expected mean squares for contrast decomposition II in equation (38).

| Factor | R<br>$a$<br>$i$ | F<br>2<br>$p$ | F<br>2<br>$q$ | R<br>$n$<br>$k$ | Expected mean square |
|---|---|---|---|---|---|
| $\tau(\beta^{FB}) \Rightarrow \tau_{i(p)}$ | 1 | 1 | 2 | $n$ | $E(MS_{A(V)}) = \sigma^2 + 2n\sigma_{\tau(\nu)}^2$ |
| $\beta^{FB} \Rightarrow \nu_p$ | $a$ | 0 | 2 | $n$ | $E(MS_V) = \sigma^2 + 2na\sum_{j=1}^{2}\nu_j^2 + 2n\sigma_{\tau(\nu)}^2$ |
| $\beta^{LR} \Rightarrow \omega_q$ | $a$ | 2 | 0 | $n$ | $E(MS_W) = \sigma^2 + 2na\sum_{j=1}^{2}\omega_j^2 + n\sigma_{\tau\omega(\nu)}^2$ |
| $\beta^D \Rightarrow (\nu\omega)_{pq}$ | $a$ | 0 | 0 | $n$ | $E(MS_{V\times W}) = \sigma^2 + na(\sum_{i=1}^{2}\sum_{j=1}^{2}(\nu\omega)_{ij}^2 + n\sigma_{\tau\omega(\nu)}^2$ |
| $[\tau \times \beta^{LR}](\beta^{FB}) \Rightarrow (\tau\omega)_{iq(p)}$ | 1 | 1 | 0 | $n$ | $E(MS_{[A\times W](V)}) = \sigma^2 + n\sigma_{\tau\omega(\nu)}^2$ |
| $\varepsilon_{(ipq)k}$ | 1 | 1 | 1 | 1 | $E(MS_E) = \sigma^2$ |

*Table 15*. Analysis of variance table for contrast decomposition I of equation (38).

| Source of variation | Sum of squares | Mean square | $F$-test |
|---|---|---|---|
| Random factor $A(B^{FB}) \Rightarrow A(V)$ | $SS_{A(V)} = SS_{A(V)} = SS_A + SS_{A\times V}$ | $MS_{A(V)} = \frac{SS_{A(V)}}{2(a-1)}$ | $F_0^{A(V)} = \frac{MS_{A(V)}}{MS_E}$ |
| Fixed factor $B^{FB} \Rightarrow V$ | $SS_V = SS_V = \frac{na}{4}[\bar{y}_{.1.} + \bar{y}_{.4.} - \bar{y}_{.2.} - \bar{y}_{.3.}]^2$ | $MS_V = SS_V$ | $F_0^V = \frac{MS_V}{MS_{A(V)}}$ |
| Fixed factor $B^{LR} \Rightarrow W$ | $SS_W = SS_W = \frac{na}{4}[\bar{y}_{.1.} + \bar{y}_{.2.} - \bar{y}_{.3.} - \bar{y}_{.4.}]^2$ | $MS_W = SS_W$ | $F_0^W = \frac{MS_W}{MS_{[A\times W](V)}}$ |
| Fixed factor $B^D \Rightarrow V \times W$ | $SS_{V\times W} = SS_{V\times W} = \frac{na}{4}[\bar{y}_{.1.} + \bar{y}_{.3.} - \bar{y}_{.2.} + \bar{y}_{.4.}]^2$ | $MS_{V\times W} = SS_{V\times W}$ | $F_0^{V\times W} = \frac{MS_{V\times W}}{MS_{[A\times W](V)}}$ |
| Random factor $[A \times B^{LR}](B^{FB}) \Rightarrow A \times W(V)$ | $SS_{[A\times W](V)} = SS_{A\times W(V)} = SS_{A\times W} + SS_{A\times V\times W}$ | $MS_{[A\times W](V)} = \frac{SS_{[A\times W](V)}}{2(a-1)}$ | $F_0^{[A\times W](V)} = \frac{MS_{[A\times W](V)}}{MS_E}$ |
| Error $E$ | $SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{ij.})^2$ | $E(MS_E) = \frac{SS_E}{4a(n-1)}$ | |
| Total | $SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} \times (y_{ijk} - \bar{y}_{...})^2$ | | |

For analysis of the nested contrast variance components of decomposition I in equation (29), the expected mean squares table is shown in Table 12, where the effect of each contrast within Factor run is random. A summary of the analysis of variances is given in Table 13.

For analysis of the nested contrast variance components of decomposition II in equation (38), the expected mean squares table is shown in Table 14, where the effect of Factor run nested by each contrast is random, and each contrast effect is fixed. A corresponding summary of the analysis of variances is given in Table 15.

## Acknowledgments

The authors appreciate the referees for their valuable comments and suggestions. In this research, Dr. J. Jin and Mr. H. Guo are partially supported by the National Science Foundation Grant: CAREER AWARD NSF DMI-0133942, and Society of Manufacturing Engineers—Education Foundation Grant: Research Initiation Award 02005.

## References

Jin, J. and Shi, J., "Diagnostic Feature Extraction from Stamping Tonnage Signals Based on Design of Experiments," *ASME Transactions, Journal of Manufacturing Science and Engineering*, Vol. 122, No. 2, pp. 360–369 (2002).

Montegomery, D. C., *Design and Analysis of Experiments*, John Wiley, New York (1997).

Nomikos, P. and MacGregor, J., "Multivariate SPC Charts for Monitoring Batch Processes," *Technometrics*, Vol. 37, pp. 41–59 (1995).

Palm, A. C., "Some Aspects of Sampling for Control Charts," *ASQC Statistics Division Newsletter*, Vol. 12, pp. 20–23 (1992).

Porter, L. J. and Caulcutt, R., "Control Charts Design: A Review of Standard Practice," *Quality and Reliability Engineering International*, Vol. 8, pp. 113–122 (1992).

Roes, K. and Does, R., "Shewhart-Type Charts in Nonstandard Situations," *Technometrics*, Vol. 37, pp. 15–40 (1995).

Runger, G. and Fowler, J. W., "Run-to-Run Control Charts with Contrast," *Quality and Reliability Engineering International*, Vol. 14, pp. 261–272 (1998).

Wetherill, G. B. and Brown, D. W., *Statistical Process Control Theory and Practice*, Chapman and Hall, London (1991).

Wheeler, D. J. and Chambers, D. S., *Understanding Statistical Process Control*, SPC Press, Knoxville, TN (1992).

Woodall, W. H. and Thomas, E. V., "Statistical Process Control With Several Components of Common Cause Variability," *IIE Transaction*, Vol. 27, pp. 757–764 (1995).

Wurl, R., Albin, S., and Shiffer, I., "Multivariate Monitoring of Batch Process Startup," *Quality and Reliability Engineering International*, Vol. 17, pp. 269–278 (2001).

Yashchin, E., "Monitoring Variance Components," *Technometrics*, Vol. 36, pp. 379–393 (1994).