Title: Analysis of Variation Transmission in Manufacturing Processes--Part II. By: Lawless, J. F., Mackay, R. J., Journal of Quality Technology, 00224065, Apr99, Vol. 31, Issue 2

Database: Business Source Complete

Analysis of Variation Transmission in Manufacturing Processes--Part I

AUTHOR:J. F. LAWLESS and R. J. MACKAY, J. A. ROBINSON

TITLE:Analysis of Variation Transmission in Manufacturing Processes--Part I

SOURCE:Journal of Quality Technology 31 no2 131-42 Ap '99

ABSTRACT

This article discusses methods of studying variation in characteristics of products that move through a multistage manufacturing process by tracking and measuring individual parts as they pass through multiple stages. We suggest simple regression and analysis of variance tools that may be used to study the amount of variation attributable to different stages of the process and to identify opportunities for variation reduction.

Key Words: Autoregressive Models, Variation Reduction.

INTRODUCTION

A fundamental strategy to reduce variation in manufacturing processes is to first identify the sources of the variation and then to take remedial action. The identification of variation sources is useful both for improving the current process and for designing more robust future processes. In processes consisting of discrete stages, there may be certain stages at which considerable variation originates and other stages that effectively absorb variation introduced upstream (i.e., at previous stages) in the process. Therefore, in order to target variation reduction activities it is important to understand how variation is added and transmitted across the stages of a process. This article discusses methodology for doing this and, in particular, for attributing the variation in key product characteristics to the various stages in a process.

We deal with situations where the same quality characteristic may be measured at the different stages of the process. A pair of examples taken from automobile manufacturing will introduce and later illustrate the main ideas. Following a description of the examples, the second section presents formal models for variation transmission, and the third section discusses the examples further. The fourth section discusses the effect of measurement error, and the fifth concludes with some remarks on extensions to the methodology. Part II of this article provides several important extensions.

EXAMPLE 1: CRANKSHAFT MACHINING

We consider the final two stages in the machining of automobile crankshafts (see Figure 1). Crankshaft journals are ground by one of four grinders and then passed to a lapper which removes some additional metal (approximately 25 × 10[sup-5] inches) to improve surface finish. The key characteristic we focus on here is a particular journal diameter, Y, measured from nominal, in units of 10[sup-5] inches. The objective is to reduce variation in Y.

Suppose the journal diameter on a crankshaft is measured before and after being processed by the lapper. We denote these measurements as Y[sub1] and Y[sub2], respectively. In addition, we define a covariate, z, such that z = j if a part is ground by grinder j.

We may partition variation in Y[sub2] into components by using the conditional variance formula (e.g., see Mood, Graybill, and Boes (1974)) as

Var(Y[sub2]) = Var [E(Y[sub2] | Y[sub1], z)] + E [Var(Y[sub2] | Y[sub1], z)]. (1)

�� The expectations on the right side of Equation (1) are with respect to Y[sub1] and z. Often, we can model Y[sub2] by a simple linear regression model in Y[sub1]; thus,

E(Y[sub2] | Y[sub1], z = j) = alpha + betaY[sub1].

Let sigma[sup2[sub[subA] = Var(Y[sub2] | Y[sub1], z). Defining sigma[sup2[sub[subi] = Var(Y[subi]) for i = 1, 2, we then obtain from Equation (1) that

sigma[sup2[sub[sub2] = beta[sup2]sigma[sup2[sub[sub1] + sigma[sup2[sub[subA]. (2)

We ignore for now the possible effects of measurement error. The first term on the right hand side of Equation (2) represents variation that arises upstream (i.e., at or before the grinding operation) and is transmitted through the lapper. The second term represents variation added by the lapping operation. If beta were close to zero, then the lapper would effectively screen out variation from upstream sources, represented by sigma[sup2[sub[sub1] = Var(Y[sub1]).

Note that, if desired, sigma[sup2[sub[sub1] can be partitioned into within- and between-grinder components. Assume that each of the four grinders processes 1/4 of the crankshafts and that each grinder is independent of the others; thus,

sigma[sup2[sub[sub1] = [Graphic Character Omitted]1/4sigma[sup2[sub[sub1j] + [Graphic Character Omitted]1/4(μ[sub1j] - μ[sub1])[sup2], (3)

where μ[sub1j] = E(Y[sub1] | z = j), sigma[sup2[sub[sub1j] = Var(Y[sub1] | z = j), and μ[sub1] = Sigma[sup4[sub[subj=1] μ[sub1j]/4. In the third section of the paper we show how these models may be used to diagnose the sources of variation in journal diameters.

EXAMPLE 2: HOOD FITS

This example is taken from the assembly of automobiles. We consider four operations widely spaced in the assembly process that relate to the installation of hoods. The four operations are: (1) install or "hang" the hood, (2) paint the hood (and the rest of the car), (3) install hardware such as the hood latch, and (4) adjust or "finesse" the hood for better fit. These operations will simply be denoted as HANG, PAINT, HARDWARE, and FINESSE. There can be physical movement of the hood or adjacent fenders at all of these operations, including PAINT, where thermal effects from the bake ovens are possible.

The quality characteristic is flushness of the hood to the surrounding fenders. We will consider four measurements: two along each side of the hood, one near the front, and one near the rear. By convention, a plus sign on the measurement indicates that the hood is high relative to the fender, and a negative sign indicates that the hood is low. The nominal condition or target is zero deviation at all locations, corresponding to a perfectly flush hood.

We are interested in changes in mean levels from one operation to the next since such changes will affect the tolerances at upstream locations required to attain the desired results at the end of the line. If mean changes were the only concern, then one could simply sample cars (not necessarily the same ones) following each of the four operations. In addition to mean changes, however, we are also interested in determining where variation is added to the process and the degree to which it is transmitted to the final flushness measurement. For this type of analysis, it is necessary to "track" vehicles through the process, taking measurements on the same vehicles at each operation, in order to estimate the statistical relationships from one operation to the next. This requires following a sample of cars and making sure that each car is met and measured at the designated operations. The measurements in the study discussed in the paper's third section were taken with a special hand-held tool. Variation due to the measurement system is a concern that we will discuss in more detail later.

As in the first example, let Y denote the measure of interest, in this case the flushness deviation in millimeters from nominal at one of the four locations. We have four operations or stages here, so Y[subi] will denote the flushness after the ith operation: i = 1, ..., 4 (corresponding to HANG, PAINT, HARDWARE, and FINESSE). Ignoring covariates other than the corresponding flushness measurement at the previous operation, the linear model discussed in the next section assumes that Y[subi] depends on Y[sub1], ..., Y[subi-1] only through Y[subi-1], with

E(Y[subi] | Y[subi-1]) = alpha[subi] + beta[subi]Y[subi-1], Var(Y[subi] | Y[subi-1]) = sigma[sup2[sub[subiA].

This gives

sigma[sup2[sub[subi] = beta[sup2[sub[subi]sigma[sup2[sub[subi-1] + sigma[sup2[sub[subiA]

for i = 2,3,4, where sigma[sup2[sub[subi] denotes the variance of Y[subi]. Here, sigma[sup2[sub[subiA] represents the variation added at the ith operation. The amount of variation after operation i that is transmitted from the previous operation is beta[sup2[sub[subi]sigma[sup2[sub[subi-1]. Hence, the regression coefficient, beta[subi], measures the degree of variation transmission.

A VARIATION TRANSMISSION MODEL

We assume that a characteristic Y is measured at each of k process stages and denote the measurement at stage i as Y[subi]. The following simple model is adequate in many situations:

Y[sub1] ˜ N (μ[sub1], sigma[sup2[sub[sub1]) (4)

Y[subi] = alpha[subi] + beta[subi]Y[subi-1] + e[subi], i = 2, ..., k, (5)

where the e[subi]'s are independent with e[subi] ˜ N (0, sigma[sup2[sub[subiA]). More specifically, this is a first order autoregressive model in which the distribution of Y[subi] given Y[subi-1], ..., Y[sub1] depends only on Y[subi-1]. This assumption can often be justified in the context of a sequential manufacturing process. The process at stage i only "knows" what is presented to it from stage i - 1 in the form of work in process. Its only "memory" of previous stages is through that work in process, which must pass through stage i - 1. The process is assumed stable in the sense that the model given by Equations (4) and (5) is valid over time.

Let us denote E(Y[subi]) by μ[subi] and Var(Y[subi]) by sigma[sup2[sub[subi]. It follows from Equations (4) and (5) that for i = 2, ..., k,

μ[subi] = alpha[subi] + beta[subi]μ[subi-1]

sigma[sup2[sub[subi] = beta[sup2[sub[subi]sigma[sup2[sub[subi-1] + sigma[sup2[sub[subiA]. (6)

Equation (6) corresponds to Equation (1). The first term on the right hand side of Equation (6) represents variation transmitted to Y[subi] from stage i - 1, and the second term represents variation added at stage i. As such, they indicate possibilities for reducing variation in Y[subi] by reducing sigma[subiA], reducing sigma[subi-1], or making beta[subi] closer to zero.

We are interested in making sigma[sup2[sub[subk] small, since it represents variation in Y at the final stage. We can consider sigma[sup2[sub[subk], in terms of variation added at stage k plus variation transmitted from stage k - 1, via Equation (6). In addition, by using Equation (6) recursively we get

sigma[sup2[sub[subk] = beta[sup2[sub[subk]beta[sup2[sub[subk-1]sigma[sup2[sub[subk-2] + beta[sup2[sub[subk]sigma[sup2[sub[subk-1,A] + sigma[sup2[sub[subkA] (7)

and, continuing back to stage 1 (and writing sigma[sup2[sub[sub1] = sigma[sup2[sub[sub1A]),

sigma[sup2[sub[subk] = (beta[subk]beta[subk-1] ... beta[sub2])[sup2] sigma[sup2[sub[sub1A] + (beta[subk]beta[subk-1] ... beta[sub3])[sup2] sigma[sup2[sub[sub2A] + ... + beta[sup2[sub[subk]sigma[sup2[sub[subk-1,A] + sigma[sup2[sub[subkA]. (8)

This decomposes the variation in Y[subk] into components attributable to each stage i = 1, ..., k and may be used to suggest the stages that might need improvements in order to reduce sigma[sup2[sub[subk]. The utility of Equation (8) is of course dependent on the validity of the autoregressive model of Equations (4) and (5), especially when there are several stages in the process. In the final section we discuss this point further and make suggestions concerning the use of Equation (8). Part II of this paper provides additional procedures for assessing the adequacy of the model in Equations (4) and (5).

ESTIMATION OF VARIATION COMPONENTS

Suppose that a study is carried out in which (Y[sub1], ..., Y[subk]) are measured on a representative set of n parts or units. That is, we assume that a random sample ((y[subj1], ..., y[subjk]), j = 1, ..., n) of measurements from the model in Equations (4) and (5) is available. In that case, it is easy to estimate the parameters μ[sub1], sigma[sub1] and alpha[subi], beta[subi], and sigma[subiA] (i = 2, ..., k): the maximum likelihood estimates (MLE's) are

[Formula Omitted]

where y[subi] = Sigma[supn[sub[subj=1] y[subji]/n, S[subii] = Sigma[supn[sub[subj=1] (y[subji] - y[subi])[sup2]/n, and S[subi-1,i] = Sigma[supn[sub[subj=1] (y[subj,i-1] - y[subi-1]) (y[subji] - y[subi])/n. The estimated version of Equation (6) is then

^sigma[sup2[sub[subi] = ^beta[sup2[sub[subi]^sigma[sup2[sub[subi-1] + ^sigma[sup2[sub[subiA], (9)

where ^sigma[sup2[sub[subi] = S[subii]. The MLE's ^sigma[sup2[sub[subi] and ^sigma[sup2[sub[subiA] involve the divisor n, so inserting estimates into Equations (6), (7), or (8) gives an exact partition of the total observed variation.

EXTENSION TO INCLUDE COVARIATES

Covariates might be introduced into the model in Equations (4) and (5) for several reasons. Suppose that at stage i, an input covariate z[subi] is measured that perhaps contributes to the added variation at that stage. The model can be modified to

Y[subi] = alpha[subi] + beta[subi]Y[subi-1] + gamma[subi]z[subi] + e*[subi], where if gamma[subi] is relatively large, the standard deviation of e*[subi] should be much less than that of e[subi]. In this instance, the added variation at stage i can be reduced by reducing variation in z[subi].

A second application of covariates is to model the effect of parallel processing streams. Suppose that at stage i there are two parallel operations, a and b, so that a part can be processed by one operation or the other. Consider the model

Y[subi] = z[subi](alpha[supa[sub[subi] + beta[supa[sub[subi]Y[supa[sub[subi-1] + e[supa[sub[subi]) + (1 - z[subi])(alpha[supb[sub[subi] + beta[supb[sub[subi]Y[supb[sub[subi-1] + e[supb[sub[subi]),

where z[subi] indicates the processing stream (z[subi] = 0, 1 for b and a) and e[supa[sub[subi], e[supb[sub[subi] have standard deviations sigma[supa[sub[subiA], sigma[supb[sub[subiA], respectively. In this model, the transmitted and added variation may be stream dependent. We can write

Var(Y[subi]) = E(Var(Y[subi] | z[subi])) + Var(E(Y[subi] | z[subi])). (10)

The second term on the right side of Equation (10) is due to the variation in targeting of the two streams. If this term is found to be large, then it can be reduced by ensuring the two streams have the same mean value, given Y[subi-1]. The first term can be partitioned as usual for each z[subi] into added and transmitted components. Aside from discussing Example 1 further in the next section, we do not consider covariates in detail, but it is clear that they can be employed usefully.

EXAMPLES

EXAMPLE 1. CRANKSHAFT MACHINING

We illustrate the ideas of the preceding section by considering the two examples introduced in the first section. Over a period of several days, n = 96 crankshafts were selected from production, and their journal diameters measured before and after passing through the lapper. Twenty-four crankshafts came from each of Grinders 1, 2, 3, and 4 (see Figure 2). Y[sub1] is the diameter before lapping, and Y[sub2] is the diameter after lapping, measured from nominal, in units of 10[sup-5] inches. The line for each plot is a least squares line, discussed below.

The scatter plots in Figure 2 are reasonably consistent with the model from the first section, where z (= 1, 2, 3, or 4) indicates which grinder processed the crankshaft and

E(Y[sub2] | Y[sub1], z) = alpha + betaY[sub1], Var(Y[sub2] | Y[sub1], z) = sigma[sup2[sub[sub2A]. (11)

The fit of the model is discussed briefly later. Let Y[sub1jz] and Y[sub2jz] denote the diameters for crankshaft j from grinder z (z = 1, 2, 3, 4; j = 1, ..., 24) before and after lapping. Define y[sub1] = Sigma[subj] Sigma[subz] y[sub1jz]/96, y[sub2] = Sigma[subj] Sigma[subz] y[sub2jz]/96. Then we have the normal distribution MLE's:

^sigma[sup2[sub[sub2] = [Graphic Character Omitted][Graphic Character Omitted](y[sub2jz] - y[sub2])[sup2]/96,

^sigma[sup2[sub[sub1] = [Graphic Character Omitted][Graphic Character Omitted](y[sub1jz] - y[sub1])[sup2]/96,

^beta = S[sub12]/^sigma[sup2[sub[sub1], and

^sigma[sup2[sub[sub2A] = ^sigma[sup2[sub[sub2] - ^beta[sup2]^sigma[sup2[sub[sub1],

where S[sub12] = Sigma[subj]Sigma[subz](y[sub1jz] - y[sub1])(y[sub2jz] - y[sub2])/96. These estimates lead to the partition of variation in Equation (2) as

^sigma[sup2[sub[sub2] = ^beta[sup2]^sigma[sup2[sub[sub1] + ^sigma[sup2[sub[sub2A],

which here gives the numerical values

11.68[sup2] = (1.026[sup2]) (10.68[sup2]) + 4.04[sup2]. (12)

The partition in Equation (12) indicates that little of the variation in the final diameter Y[sub2] is added by the lapper: most of it is due to variation in the pre-lapper diameter Y[sub1] that is then transmitted through the lapper. If we break ^sigma[sup2[sub[sub1] into within- and between-grinder components, then we find, analogous to Equation (3), that 10.68[sup2] = 9.22[sup2] + 5.39[sup2]. Most of the variation in Y[sub1] is therefore due to variation in the crankshafts that come out of each grinder. Note that if we had not measured the diameter Y[sub1] of each crankshaft before lapping but had observed which grinder the crankshaft came from, then we would be able to determine that not much of the variation in Y[sub2] was due to differences between crankshafts processed by the four grinders. However, we would not know that most of the variation was transmitted from upstream, as opposed to being added by the lapper.

Figure 2 shows the least squares (maximum likelihood) line y[sub2] = y[sub2] + ^beta(y[sub1] - y[sub1]) superimposed on the scatter plot for each grinder. The data are reasonably consistent with the model in Equation (11); there is evidence of very mild departures, especially for Grinder 3. However, if we fit models that accommodate these small departures, (e.g., by allowing nonlinear regression curves or separate regression lines for each grinder), then the added features are not statistically significant, the partition of variation changes very little, and the overall conclusions are unchanged. Hence, we prefer to retain Equation (11) as a useful and quite accurate summary of variation in the data. The main conclusion from the analysis is that efforts should be concentrated on reducing within-grinder variation.

EXAMPLE 2. HOOD FITS

This section contains an analysis of data from the hood fits example, based on a study involving n = 19 cars. The discussion will center around the set of graphs given in Figures 3-7. The sequence or line charts in Figure 3 contain plots of the flushness measured at each of the four plant operations for each of the four measurement locations. Each connected set of points represents one car. The most obvious observation is that the hood is moving up relative to the fenders at the rear (i.e., the measurements are becoming less negative) during processing. The statistics indicate that the movement is approximately 1.5 mm on each side, most of it accounted for at FINESSE. The end result is that near nominal fits are attained on average on the left side, but hoods remain 1.66 mm low on the right. While seemingly trivial, we have found this type of finding to be extremely important in practice. For example, here it indicates that to remove the FINESSE operation, it would be necessary to change tolerances so that hoods are installed higher at the rear than is the current practice.

As mentioned earlier, it is not necessary to track vehicles to determine this type of change in mean. There are, however, more complex questions that require tracking vehicles: What are the sources of the variation? Is it being passed through the system, or is it being added at a few key operations? The line charts provide a glimpse of the answers to these questions. Parallel lines between stages indicate transmission of variation. With parallel lines the mean may change, but the variation at the end is the same as at the beginning. Lines that splay out indicate a magnification of variation, while converging or criss-crossing lines indicate some degree of adjustment. When there is an adjustment at an operation, large values are made smaller, and small values are made larger. With these general guidelines in mind, it seems clear in this example that there is a great deal of transmitted variation at the rear of the hood and a lesser amount at the front. However, this can be displayed more clearly with scatter plots.

Figure 4 contains scatter plots of the deviations following the PAINT operation against the corresponding deviations following the HANG operation. Three of the four regression slopes, including the two corresponding to the rear of the hood, are near one, indicating a high degree of transmitted variation. The residual standard deviations can be interpreted as indicating the amount of variation added at PAINT. Figure 5 displays the scatter plots for the next pair of consecutive operations, PAINT and HARDWARE. The results are similar except the added variation (scatter) is larger at the front than before. Figure 6 shows the scatter plots for the final two operations, HARDWARE and FINESSE. Again, the regression slopes are near one at the rear. These scatter plots bear out the general conclusion of a high degree of transmitted variation at the rear of the hood. This means that variation that exists after the hood is hung is being passed on to the end of the line. We conclude that end-of-line variation at the rear of the hood could be reduced if variation at HANG were reduced.

The front of the hood behaves differently. There is strong transmission of variation through HARDWARE on the left side, but it is much weaker on the right side (Figures 4 and 5). Also, there is essentially no transmission of variation through FINESSE on the left side (Figure 6). It is operating as a pure adjuster--the output being completely independent of the input. For all operations, the added variation, as measured by the residual standard deviation, is greater at the front than at the rear.

It helps to quantify the results by showing the variance decomposition in Equation (8) into components corresponding to individual operations. A stacked bar chart is a convenient way to display the results. The height of each bar is the total variation immediately following that operation. The components of variation make up the elements of the stack. The convention followed here is to place the component corresponding to the first operation at the bottom and then proceed upward in order, so that the top component corresponds to the variation "added" at that operation. One should bear in mind that the chart portrays variances, but that the effects on the standard deviation of Y_k are of primary interest.

Figure 7 displays the bar charts for the hood data. It confirms and quantifies the previous conclusions. Variation is being passed through at the rear, but there is no corresponding transmission of variation at the front beyond HARDWARE. At the front, nearly all of the variation following FINESSE is due to that operation (left = 100%, right = 80%). This means that to benefit in an overall way from reducing variation at HANG, we would need to eliminate (or change) the FINESSE operation. As noted previously, targets would have to be adjusted to change mean levels.

We can focus on the third bars in Figure 7 to understand the impact of removing the FINESSE operation. We might conclude that to reduce variation below current levels at the front, we would need to improve both the HANG and HARDWARE operations. Obviously, in practice this would have to be discussed in terms of the engineering understanding and impact on the process.

Model checks do not show any serious problems. In the next section we discuss the effect of measurement error on the type of analysis used here. Measurement error is considered negligible for the hood data, but the methods in Part II allow an analysis that incorporates it. We find that the conclusions are unchanged.

IMPORTANT POINTS RELATED TO THE METHODOLOGY

The analysis described above depends on the approximate validity of the autoregressive model in Equations (4) and (5). Diagnostic checks based on regression fits and residuals should therefore be part of the analysis. This is discussed in some detail in Part II. In this section, we mention three important points related to the methodology used in the current paper.

The first point concerns measurement error. We have assumed that Y_i is measured exactly, but this is typically not the case. To see the effect of measurement error, suppose that, instead of Y_i, we observe

$X_i = Y_i + \epsilon_i$,

where the ϵ_i's (i = 1, 2, ..., k) are independent $N(0, \sigma^2_{\epsilon_i})$ variables and are independent of the Y_i's. If we proceed with our analysis by treating the X_i's as the Y_i's, then we are in effect using the incorrect model, and it may easily be shown (see, e.g., Fuller (1987)) that in large samples

$\hat{\sigma}^2_i$ [Graphic Character Omitted] $\sigma^2_i + \sigma^2_{\epsilon_i} = Var(X_i)$ (13)

$\hat{\beta}_i$ [Graphic Character Omitted] $(\beta_i \sigma^2_{i-1}) / (\sigma^2_{i-1} + \sigma^2_{\epsilon_i}) = Cov(X_{i-1}, X_i) / Var(X_{i-1})$. (14)

Thus, $\hat{\beta}_i$ tends to underestimate β_i; this is the well known measurement error attenuation effect.

We need to consider Equations (6) and (9) along with Equations (13) and (14) in order to see the effects of the measurement error on the variance decomposition and hence on our assessment of the possibilities for reducing variation. Table 1 summarizes the effects of ignoring measurement error on the estimates of the three components of variation at a given stage: variation transmitted from the previous stage, variation added at the stage, and measurement variation. Transmitted variation is underestimated by the same proportional factor affecting the regression slope. The estimate of added variation is biased upward by the measurement error, as expected. But there is an additional positive bias equal in magnitude to the negative bias for transmitted variation.

If $\sigma_{\epsilon_{i-1}}$ is sufficiently small so that $\sigma^2_{i-1} / (\sigma^2_{i-1} + \sigma^2_{\epsilon_{i-1}})$ exceeds, say, 0.9, then the effects of measurement error may be deemed relatively unimportant and can be ignored. However, if there are several stages in the process and we use Equation (8) to attribute variation to the different stages, then for earlier stages the underestimation is much more severe because of the presence of terms involving products of the β_i's. For example, the proportional bias factor associated with the amount of variance transmitted from stage s, or earlier, to stage k, is $\Pi^{k-1}_{i=s+1} \sigma^2_i / (\sigma^2_i + \sigma^2_{\epsilon_i})$, which becomes smaller as s becomes smaller, indicating a greater degree of underestimation for "far upstream" stages. Consequently, Equation (8) should be used with caution if k is large. If the measurement error variances, $\sigma^2_{\epsilon_i}$, are known, then it is possible to make adjustments to the analysis, the simplest approach being to employ Equations (13) and (14). Part II of this article provides a full discussion of the measurement error and shows how to extend these methods to deal with it.

A second caveat concerns the validity of the model in Equations (4) and (5) more generally. Occasionally there may be situations where, even if the measurement error is insignificant, the distribution of Y_i may depend on not only Y_{i-1} but also on earlier measurements. In this case, the variance decomposition in Equation (6) is still valid when the measurements are approximately jointly normally distributed, and it provides insight into the transfer of variation from stage i - 1 to stage i. However, multistage Equations (7) and (8) do not usually provide much insight, and it may even be difficult to assess the effect of changes to stage i - 1 on the stage i measurements, given the complex relationship between the different measurements. Diagnostic checks on the dependence of Y_i on only Y_{i-1} may be made by regressing Y_i on Y_{i-1} and upstream measurements. Part II discusses further checks of the autoregressive model. We recommend that when checks indicate that the first order autoregressive model is unsatisfactory, the joint distribution of (Y_1, ..., Y_k) be examined carefully.

Our third point concerns the fact that the analysis in the current paper is based entirely on point estimates and graphs. Although this informal analysis is often satisfactory, uncertainty due to sampling variation should be considered. Confidence intervals or tests for variance components or regression coefficients may be obtained. Part II shows how to do this and considers the hood flushness data as an example. The confidence interval procedures also provide information on the number of parts needed for a study to be effective.

CONCLUSION

This article has discussed how to study variation in key characteristics as discrete parts move through a multistage production process by tracking and measuring the characteristics of individual parts through the stages of the process. In many situations we can determine how much variation is added at different stages and how much of that variation is transmitted downstream. To use the methodology, it is necessary to identify a characteristic for study, have a capable process for measuring the characteristic, and have the ability to follow individual parts through the process. It is important that measurement error be small and that the first order autoregressive model described in the second section provides a reasonably accurate representation of the measurements. Checks of these assumptions should therefore be included as part of the application of these methods.

Whether measurement error is suitably "small" can be assessed through the results in the fourth section. If $\sigma_{\epsilon_i}/\sigma_i$ is less than, say, .1, then the measurement error can be ignored when the number of stages is only two or three. If desired, estimates of β_i's and σ_i's and corresponding partitions of variance can be adjusted using Equations (13) and (14). In Part II, more formal methods of dealing with measurement error are given, and methods for obtaining confidence limits and for model checking are presented.

For most processes there will be several characteristics of interest; for example, the hood fitting process discussed earlier had many key measurements, four of which were discussed. The characteristics can be studied individually, as in the second example in the third section. However, this does not take account of the way that the multiple characteristics interact, and in some cases this may lead to important opportunities for variation reduction being missed. Correlated measurements may also make separate first order autoregressive models for single characteristics inappropriate. Methodology for variation analysis based on multivariate autoregressive models is discussed by Fong and Lawless (1998), but further work in this area is needed.

The discussion in this article has emphasized methods for studying the transmission of variation across stages in a process in order to identify opportunities for reducing variation. We have not specifically discussed targeting of the characteristics, but this is easily done. In addition, we have considered situations where the same characteristics exist and may be measured at different process stages. Similar ideas may be used more generally to study the relationship between upstream variables $z_1, \ldots, z_k$ and one or more quality characteristics Y on a finished part. The idea is to assess variation in $(z_1, \ldots, z_k)$ and how it affects Y. In general, the variables may be of various types (categorical, continuous, etc.) and relationships may be non-normal and non-linear; thus, rather different models than the ones used in this paper may be needed. These ideas will be considered elsewhere, but some analogies with the work of Taguchi and others (e.g., Taguchi, Elsayed, and Hsiang (1989)) on the analysis of variation in products and systems are apparent.

ADDED MATERIAL

J. F. LAWLESS and R. J. MACKAY

University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

J. A. ROBINSON

General Motors Global R&D Operations, Warren, MI 48090

Dr. Lawless is a Professor in the Department of Statistics and Actuarial Science. He holds the General Motors Canada-NSERC Industrial Research Chairman Quality and Productivity. His email address is jlawless@uwaterloo.ca.

Dr. MacKay is a Professor in the Department of Statistics and Actuarial Science.

Dr. Robinson is Principal Research Scientist in the Manufacturing and Design Department.

ACKNOWLEDGMENTS

This research was supported by General Motors and by grants to the first two authors from the Natural Sciences and Engineering Research Council of Canada.

TABLE 1. Bias of Components of $\mathrm{Var}(Y_i)$ When Ignoring Measurement Variation

Source	Actual	Estimated
Transmitted	$\beta_i^2\sigma_{i-1}^2$	$\beta_i^2\sigma_{i-1}^2$
Added	σ_{iA}^2	$\sigma_{iA}^2 + \sigma_{\epsilon_i}^2 + \beta_i^2\sigma_{i-1}^2(1 - R_i)$
Measurement	$\sigma_{\epsilon_i}^2$	0

Where $R_i = \sigma_{i-1}^2/(\sigma_{i-1}^2 + \sigma_{\epsilon_i}^2)$

Source	Bias
Transmitted	$-\beta_i^2\sigma_{i-1}^2(1 - R_i)$
Added	$\sigma_{\epsilon_i}^2 + \beta_i^2\sigma_{i-1}^2(1 - R_i)$
Measurement	$-\sigma_{\epsilon}^2$

Where $R_i = \sigma_{i-1}^2/(\Sigma_{i-1}^2 + \sigma_{\epsilon_i}^2)$

FIGURE 1. Grinder and Lapper Stages in a Crankshaft Machining Process.

FIGURE 2. Crankshaft Diameters Before and After Lapping.

FIGURE 3. Flushness at Four Plant Operations.

FIGURE 4. Flushness After HANG and PAINT Operations.

FIGURE 5. Flushness After PAINT and HARDWARE Operations.

FIGURE 6. Flushness After HARDWARE and FINESSE Operations.

FIGURE 7. Flushness Variation by Source at Four Plant Operations.

REFERENCES

FONG, D. Y. T. and LAWLESS, J. F. (1998). "The Analysis of Process Variation Transmission with Multivariate Measurements". Statistica Sinica 8, pp. 151-164.

FULLER, W. A. (1987). Measurement Error Models. John. Wiley & Sons, New York, NY.

MOOD, A. M.; GRAYBILL, F. A.; and BOES, D. C. (1974). Introduction to the Theory of Statistics. McGraw-Hill, New York, NY.

TAGUCHI, G.; ELSAYED, E.; and HSIANG, T. (1989). Quality Engineering in Production Systems. McGraw-Hill, New York, NY.

Source: Journal of Quality Technology, Apr99, Vol. 31 Issue 2, p143, 12p
Item: 1733240