

THE ANALYSIS OF PROCESS VARIATION TRANSMISSION WITH MULTIVARIATE MEASUREMENTS

Daniel Y. T. Fong and J. F. Lawless

University of Waterloo

Abstract: This article discusses methods of estimating the variation in product quality characteristics measured at several stages in a manufacturing process. By determining which stages contribute most to variation one can focus variation reduction activities more effectively. A multivariate normal Markov process is used to model the variation in characteristics. Methods that deal with measurement error and missing data are introduced through a state space formulation.

Key words and phrases: E-M algorithm, Kalman filter, measurement error, missing data.

1. Introduction

In order to reduce variation in manufacturing processes consisting of several discrete stages it is often worthwhile to study the variation that is added at different stages, and whether that variation is transmitted downstream to subsequent stages. In particular, there may be certain stages where considerable variation originates, and other stages that filter out variation introduced upstream. By understanding how variation is added and transmitted across the stages of a process we can decide where to concentrate variation reduction efforts.

For illustration we consider a pair of examples taken from automobile manufacturing.

Example 1. Hood Fits

Lawless, MacKay and Robinson (1996) discuss an assembly process that is part of the installation of car hoods. There are four stages of the process, corresponding to four operations: (1) install or “hang” the hood, (2) paint the hood (and the rest of the car), (3) install hardware such as the hood latch, (4) adjust or “finesse” the hood for better fit. The quality characteristics of interest relate to the flushness of the hood to the surrounding fenders of the car. This is quantified through four flushness deviation measurements, one near the front and one near the rear on each side of the hood. A zero measurement at any location means the hood is perfectly flush, and positive and negative measurements mean it is too high and too low, respectively.

The objective is to have hoods as close to perfectly flush as possible after the final stage. By “tracking” vehicles and taking flushness measurements after each stage, it is possible to learn about the origins and transmission of variation.

Example 2. Piston Machining

Agrawal, Lawless and MacKay (1996) examined a study on the machining of pistons. These are essentially cylinders which are closed at the top and open at the bottom, where they are connected to rods. The quality characteristics of interest were four diameters, located at heights of 4 mm, 10 mm, 36.7 mm and 58.7 mm from the bottom of the piston. The diameters were measured after each of several operations in the machining process, the measurements being in millimeters, to a precision of 1 micron (10^{-3} mm). The objective is to produce pistons for which all four diameters are very close to specified values.

Lawless, MacKay and Robinson (1996) presented methods for analyzing the transmission of variation in a univariate characteristic, based on a first order autoregressive model. In order to carry out such analysis it is necessary to be able to track units (in our examples these are vehicles) through the manufacturing process so that measurements may be taken on the same unit at different stages. Lawless et al. (1996) assumed that a univariate quality characteristic y_t is measured at each of T process stages $t = 1, \dots, T$, and considered the model

$$y_1 = \mu_1 + e_1 \tag{1.1}$$

$$y_t = \alpha_t + \beta_t y_{t-1} + e_t \quad t = 2, \dots, T, \tag{1.2}$$

where $e_t \sim N(0, \sigma_{e_t}^2)$ and are independent. This first order Markov, or autoregressive AR(1) model can often be justified in manufacturing processes, and it leads to the following variation transmission formula for $\sigma_t^2 = \text{Var}(y_t)$:

$$\sigma_t^2 = \beta_t^2 \sigma_{t-1}^2 + \sigma_{e_t}^2. \tag{1.3}$$

The first term on the right side of (1.3) represents variation transmitted from stage $t - 1$ to stage t , and the second term represents variation added at stage t . Lawless et al. (1996) fitted models (1.1) and (1.2) to process data and discussed how to use (1.3) recursively to assess variation transmission across stages $t = 1, \dots, T$ of a process.

The present paper extends the techniques of Lawless et al. (1996) in several directions. First, we consider multivariate measurements, and in particular, deal with a multivariate version of (1.1) and (1.2). We refer to the model as an AR(1) model, but it should be noted that T is generally small and the model is non-stationary, unlike many applications involving AR(1) models. Second, we deal with missing data; this is important since it is often difficult to measure

all characteristics on every unit in a study that is undertaken on-line, i.e. while the manufacturing process is operating. Finally, we incorporate measurement error into the multivariate AR(1) model; this is important because, as discussed by Agrawal et al. (1996) and Lawless et al. (1996), if substantial measurement error is ignored the results of the AR(1)-based variance transmission analysis are misleading.

Section 2 of the paper introduces the multivariate AR(1) model and incorporates measurement error. Section 3 is the core of the paper and presents methodology for fitting the model to process data; this is done by using a state space formulation that leads to efficient computational procedures. Section 4 illustrates the methodology on the piston machining process, and Section 5 concludes with comments and points that deserve further study.

2. An AR(1) Variation Transmission Model

The methods in this paper are designed for use on a stable process. That is, the model (2.1)-(2.2) applies to units manufactured over time, and the parameter values in the model do not change over time. We assume that sequential measurements on a random sample of n units from the process are available. As discussed by Lawless et al. (1996) for the univariate case, we consider a (non-stationary) first order autoregressive, or AR(1), model for the $C \times 1$ vector of multivariate measurements z_{it} on unit i at stage t ($t = 1, \dots, T$; $i = 1, \dots, n$). This can be expressed as

$$z_{i1} = \boldsymbol{\mu}_1 + \mathbf{e}_{i1} \quad (2.1)$$

$$z_{it} = \mathbf{A}_t + \mathbf{B}_t z_{i,t-1} + \mathbf{e}_{it} \quad t = 2, \dots, T, \quad (2.2)$$

where $\mathbf{e}_{it} \sim N_c(\mathbf{0}, \Sigma_{e_t})$, $t = 1, \dots, T$; the notation $\mathbf{y} \sim N_p(\boldsymbol{\mu}, \Sigma)$ means that \mathbf{y} has a p -variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ . The dimensions of \mathbf{A}_t and \mathbf{B}_t are $C \times 1$ and $C \times C$, respectively. It is assumed that the measurements for different units are independent.

The marginal means and covariance matrices for the z_{it} 's are given by

$$E(z_{i1}) = \boldsymbol{\mu}_1, \quad E(z_{it}) = \boldsymbol{\mu}_t = \mathbf{A}_t + \mathbf{B}_t \boldsymbol{\mu}_{t-1}, \quad t = 2, \dots, T \quad (2.3)$$

$$\text{Var}(z_{i1}) = \Sigma_{e_1}, \quad \text{Var}(z_{it}) = \Sigma_t = \mathbf{B}_t \Sigma_{t-1} \mathbf{B}_t' + \Sigma_{e_t}, \quad t = 2, \dots, T. \quad (2.4)$$

In addition

$$\text{Cov}(z_{is}, z_{it}) = \Sigma_{st} = \sum_s \mathbf{B}_{s+1}' \dots \mathbf{B}_t' \quad (s < t). \quad (2.5)$$

The vector \mathbf{e}_t and its covariance matrix Σ_{e_t} represent variation added at stage t , whereas $\mathbf{B}_t \Sigma_{t-1} \mathbf{B}_t'$ represents variation transmitted from stage $t-1$; in this regard the right hand portion of (2.4) is the multivariate generalization of (1.3). The intercept \mathbf{A}_t allows the means $\boldsymbol{\mu}_t = E(z_{it})$ to vary across $t = 1, \dots, T$. In a

case like that in Example 2, for instance, a stage may reduce the diameters from the preceding stage substantially. An alternative but equivalent parameterization is $E(z_{it}|z_{i,t-1}) = \mu_t + B_t(z_{i,t-1} - \mu_{t-1})$.

In practice there may be significant measurement error, that is, variation in the process by which the $\mathbf{z}_{it} = (z_{it1}, \dots, z_{itc})'$ are measured. As discussed in Section 5, this can invalidate the methods described herein if it is ignored, so we consider it explicitly. We let \mathbf{y}_{it} represent the measurement of \mathbf{z}_{it} and assume that

$$\mathbf{y}_{it} = \mathbf{z}_{it} + \boldsymbol{\delta}_{it}, \quad t = 1, \dots, T, \quad (2.6)$$

where the $\boldsymbol{\delta}_{it}$'s are mutually independent $N_c(\mathbf{0}, \Sigma_{\delta_t})$ random vectors and are independent of the \mathbf{e}_{it} 's in (2.1) and (2.2). It should be noted that the \mathbf{y}_{it} 's do not follow an AR(1) model.

The motivation for considering the model (2.2) is to examine the sources of variation in the measurements \mathbf{z}_{iT} at the final stage. This may be done by working backwards from the final stage: (2.2) for $t = T$ indicates that the covariance matrix Σ_T may be decomposed into variation transmitted from stage $T - 1$ and variance added at stage T ,

$$\Sigma_T = B_T \Sigma_{T-1} B_T' + \Sigma_{e_T}. \quad (2.7)$$

Similarly, Σ_{T-1} may be decomposed and, working backwards, we may ascertain the contribution of the variation added at any stage t (i.e. Σ_{e_t}) to Σ_T . Multivariate covariance matrices may admittedly be hard to interpret, and it is important to relate them to the physical properties of the units under consideration. The example of Section 5 illustrates and discusses this further.

Care should be taken to assess the appropriateness of the model (2.1)-(2.2), possibly with measurement error accounted for by (2.6). Section 4 discusses model checking and Section 5 comments on the robustness of the methods to departures from the model.

3. Parameter Estimation

It is important to have estimation procedures that deal with missing data, since it is often impossible to measure all the characteristics on every unit at every stage. We therefore suppose that some arbitrary subset of the CT univariate measurements on unit i may be missing, and that observations are missing completely at random in the terminology of Rubin (1976) and Little and Rubin (1987). This means that the probability a particular set of measurements on a unit is missing does not depend on the values of the measurements for that or other units, and implies that the likelihood function may be based on the joint distribution of the measurements available for each unit.

We assume that the covariance matrices \sum_{δ_t} ($t = 1, \dots, T$) for the measurement errors are known. In practice these should be estimated from measurement studies. The set of unknown parameters then includes μ_1 , the \sum_{e_t} 's ($t = 1, \dots, T$) and the A_t 's and B_t 's ($t = 2, \dots, T$). Since the observed measurements y_{itc} ($t = 1, \dots, T; c = 1, \dots, C$) for unit i jointly follow a multivariate normal distribution of dimension CT or less, it would be possible in principle to write the mean and covariance matrix for each i in terms of the unknown parameters and to maximize the likelihood by a search algorithm. In particular, we note that, under (2.1), (2.2) and (2.6), the complete data y_{it} 's have means μ_t given by (2.3) and covariance matrices

$$\text{Var}(\mathbf{y}_{it}) = \sum_t + \sum_{\delta_t}, \quad \text{Cov}(\mathbf{y}_{is}, \mathbf{y}_{it}) = \sum_{st} \quad (s < t), \quad (3.1)$$

where \sum_t and \sum_{st} are given in (2.4) and (2.5), respectively. This brute force approach encounters matrices of large dimension if CT is large, and is computationally slow; the latter is a drawback for the use of bootstrap methods for obtaining variance estimates or confidence intervals, as described in Section 4. Consequently we express the model in state space form (e.g. Harvey (1989), Harvey and McKenzie (1984), Shumway (1988)), and utilize the $E - M$ algorithm (Dempster, Laird and Rubin (1977)) to obtain maximum likelihood estimates.

The model given by (2.1), (2.2) and (2.6) with arbitrary measurements missing at random can be expressed in the following form, where \mathbf{y}_{it} now stands for the vector of observed measurements on unit i at stage t :

$$\mathbf{z}_{it} = \mathbf{A}_t + B_t \mathbf{z}_{i,t-1} + \mathbf{e}_{it} \quad (3.2)$$

$$\mathbf{y}_{it} = H_{it} \mathbf{z}_{it} + H_{it} \boldsymbol{\delta}_{it}, \quad (3.3)$$

where $i = 1, \dots, n$; $t = 1, \dots, T$, we define $\mathbf{A}_1 = \boldsymbol{\mu}_1$, $B_1 = 0$, $\mathbf{z}_{i1} = 0$, and where H_{it} is a matrix obtained by taking the $C \times C$ identity matrix and deleting rows which correspond to missing observations on unit i at stage t .

The log likelihood function based on the observed data may be written in the form of an arbitrary constant plus

$$\ell = -\frac{1}{2} \sum_{i,t} \log |\sum_{iy} (t|t-1)| - \frac{1}{2} \sum_{i,t} (\mathbf{y}_{it} - \mathbf{y}_{i,t|t-1})' \sum_{iy}^{-1} (t|t-1) (\mathbf{y}_{it} - \mathbf{y}_{i,t|t-1}), \quad (3.4)$$

where we introduce the notation

$$\mathbf{y}_{ip|q} = E[\mathbf{y}_{ip} | \mathbf{y}_{i1}, \dots, \mathbf{y}_{iq}], \quad \sum_{iy} (p|q) = \text{Var}[\mathbf{y}_{ip} | \mathbf{y}_{i1}, \dots, \mathbf{y}_{iq}] \quad (3.5)$$

and where the range for i and t in the sum $\sum_{i,t}$ is over $i = 1, \dots, n$ and $t = 1, \dots, T$. Expression (3.4) assumes there is at least one measurement at each

stage for each unit. If all measurements at a stage t happen to be missing for unit i , then (3.4) is modified to omit terms involving $\sum_{iy}(t|t-1)$, $\sum_{iy}(t+1|t)$ and to add a term involving $\sum_{iy}(t+1|t-1)$.

The terms $y_{i,t|t-1}$ and $\sum_{iy}(t|t-1)$ needed to calculate (3.5) may be computed recursively using the following state space, or Kalman *filtering formulas*: define, following (3.5),

$$z_{ip|q} = E[z_{ip}|y_{i1}, \dots, y_{iq}], \quad \sum_{iz}(p|q) = \text{Var}[z_{ip}|y_{i1}, \dots, y_{iq}],$$

and set $z_{i0|0} = 0$, $\sum_{iz}(0|0) = 0$. Then for $t = 1, \dots, T$

$$z_{it|t-1} = A_t + B_t z_{i,t-1|t-1} \tag{3.6}$$

$$\sum_{iz}(t|t-1) = B_t \sum_{iz}(t-1|t-1) B_t' + \sum_{e_t} \tag{3.7}$$

$$y_{it|t-1} = H_{it} z_{it|t-1} \tag{3.8}$$

$$\sum_{iy}(t|t-1) = H_{it} \sum_{iz}(t|t-1) H_{it}' + H_{it} \sum_{\delta_t} H_{it}' \tag{3.9}$$

where $z_{it|t}$ and $\sum_{iz}(t|t)$ are computed via

$$P_{it} = \sum_{iz}(t|t-1) H_{it}' \sum_{iy}(t|t-1)^{-1}$$

$$z_{it|t} = z_{it|t-1} + P_{it}(y_{it} - y_{it|t-1})$$

$$\sum_{iz}(t|t) = \sum_{iz}(t|t-1) - P_{it} \sum_{iy}(t|t-1) P_{it}'.$$

Derivation of these formulas is outlined in the Appendix. These calculations involve only square matrices of dimension C or smaller.

Now that we can compute (3.4), we could maximize it by using a derivative-free procedure such as the simplex search algorithm (Nelder and Mead (1965), Press et al. (1986), Section 10.4). An attractive alternative, which also allows easier access to model-checking and to handling cases where entire stages are missing on some units, is an $E - M$ algorithm. This has been well discussed for use with missing data in normal models (e.g. Little and Rubin (1987), Chapter 8) and is adapted here to deal with both missing data and measurement error.

Referring to (3.2), we consider the “complete data” log likelihood as that based on the z_{it} 's, which may be written as an arbitrary constant plus

$$\ell_c = -\frac{nC}{2} \sum_{t=1}^T \log |\sum_{e_t}| - \frac{1}{2} \sum_{i=1}^n \sum_{t=1}^T e_{it}' \sum_{e_t}^{-1} e_{it}. \tag{3.10}$$

The model (3.2) is AR(1), and maximum likelihood estimates are easily found to be (e.g. Mardia, Kent and Bibby (1979), Chapter 6)

$$\hat{A}_1 = \bar{z}_1, \quad \hat{\sum}_{e_1} = S_{1,1}$$

$$\begin{aligned}\hat{B}_t &= S_{t,t-1}(S_{t-1,t-1})^{-1}, & \hat{A}_t &= \bar{z}_t - \hat{B}_{t-1}\bar{z}_{t-1} \\ \hat{\sum}_{e_t} &= S_{t,t} - \hat{B}_t S_{t-1,t}\end{aligned}\quad (3.11)$$

for $t = 2, \dots, T$, where

$$\bar{z}_t = \frac{1}{n} \sum_{i=1}^n z_{it}, \quad S_{u,t} = \frac{1}{n} \sum_{i=1}^n z_{iu} z'_{it} - \bar{z}_u \bar{z}'_t. \quad (3.12)$$

The M -step in the $E - M$ algorithm is given by (3.11). The E -step consists of computing the expectations of the complete data, conditional on the observed data, that are needed to compute the conditional expectation of (3.10). This may be done using the state-space *smoothing formulas* for $t = 1, \dots, T - 1$:

$$\begin{aligned}R_{it} &= \sum_{iz} (t|t) B'_{t+1} \sum_{iz} (t+1|t)^{-1} \\ z_{it|T} &= z_{it|t} + R_{it}(z_{i,t+1|T} - z_{i,t+1|t})\end{aligned}\quad (3.13)$$

$$\sum_{iz} (t|T) = \sum_{iz} (t|t) - R_{it}[\sum_{iz} (t+1|t) - \sum_{iz} (t+1|T)] R'_{it}. \quad (3.14)$$

Derivations are outlined in the Appendix. The E -step is now carried out by replacing \bar{z}_t and $z_{i,t-1} z'_{it}$ in the expressions (3.12) with (compare Little and Rubin (1987), page 143)

$$\frac{1}{n} \sum_{i=1}^n z_{it|T} \quad (3.15)$$

$$z_{i,t-1|T} z'_{it|T} + \text{Cov}(z_{i,t-1}, z_{it} | y_{i1}, \dots, y_{iT}) \quad (3.16)$$

respectively, evaluated at the most recent parameter estimates from the M -step (3.11).

In the case where there is no measurement error, $\text{Cov}(z_{i,t-1}, z_{it} | y_{i1}, \dots, y_{iT}) = \sum_{iz} (t-1|T) B'_t$. More generally, however, it must be obtained from the smoothing formula (3.14) for the augmented model

$$z_{it}^* = \begin{pmatrix} z_{it} \\ z_{i,t-1} \end{pmatrix} = \begin{pmatrix} A_t \\ 0 \end{pmatrix} + \begin{pmatrix} B_t & 0 \\ I & 0 \end{pmatrix} z_{i,t-1}^* + \begin{pmatrix} e_{it} \\ 0 \end{pmatrix},$$

where I represents an identity matrix.

The $E - M$ algorithm proceeds by alternating E and M steps until convergence is achieved. Initial estimates that can be used to start the process can be obtained by the following simple procedure: compute empirical means \bar{y}_t and cross-product matrices $S_{t,t}$ and $S_{t-1,t}$ using units with no missing measurements at stage t (for \bar{y}_t and $S_{t,t}$) and at stages $t - 1$ and t (for $S_{t-1,t}$), respectively. Then, compute the estimates

$$\tilde{\mu}_1 = \bar{y}_1 \quad \tilde{\sum}_{e_1} = S_{1,1} - \sum_{\delta_1} \quad (3.17)$$

$$\begin{aligned}\tilde{B}_t &= S_{t,t-1}(S_{t-1,t-1} - \sum_{\delta_{t-1}})^{-1} & t = 2, \dots, T \\ \tilde{A}_t &= \tilde{\mathbf{y}}_t - \tilde{B}_t \tilde{\mathbf{y}}_{t-1} & t = 2, \dots, T \\ \tilde{\sum}_{e_t} &= (S_{t,t} - \sum_{\delta_t}) - \tilde{B}_t S_{t-1,t} & t = 2, \dots, T.\end{aligned}$$

When there is no missing data, these are the estimates that would be obtained by maximum likelihood if the process had only $T = 2$ stages. Agrawal et al. (1996) study these estimates in the univariate case.

There are many $(CT + C^2(T - 1) + C(C + 1)T/2)$ parameters in the model, and we are primarily interested in components of variance as epitomized in (1.3) and (2.4). In these circumstances it does not make sense to develop estimates of the asymptotic variances and covariances of all parameter estimates. In order to assess variation in estimates and to obtain confidence intervals for quantities of interest, we use a parametric bootstrap (Efron and Tibshirani (1993)). The procedure is as follows: treating the maximum likelihood estimates as if they were the true parameter values and the H_{it} 's as given by the pattern of missingness in the original data, we generate B sets of data from the model (3.2)-(3.3). For each of the B sets of data we obtain maximum likelihood estimates $\hat{\theta}_b^*$ (where θ stands for the vector of all parameters). Estimates of functions $\psi = g(\theta)$ that are of interest are then calculated for each sample. Variance estimates for $\hat{\psi} = g(\hat{\theta})$ (where $\hat{\theta}$ is the maximum likelihood estimate from the original data) or confidence intervals for ψ may then be calculated in various standard ways (see Efron and Tibshirani (1993)).

An example of the bootstrap methods is given in Section 4.

4. An Example

We consider data on 96(= n) randomly selected pistons from the process mentioned in Example 2 of Section 1. Four (= C) diameter measurements were taken at each of 4(= T) process stages.

The model represented by (2.1), (2.2) and (2.6) was fitted. There are no missing observations here and the measurement error covariance matrix is assumed to be $\sigma_\delta^2 I_4$, where I_4 is the 4×4 identity matrix. The measurements are discrete, diameters being measured to the nearest micron ($10^{-3}mm$), and at each of the 4 locations on the piston fewer than 15 distinct values occur. Nevertheless we work with the assumed normal model, which seems to provide a reasonable picture of variation.

Models were fitted with $\sigma_\delta^2 = .04167$ microns² and also with $\sigma_\delta^2 = .1$ microns². The former corresponds to the variance of a triangular distribution on $(-.5, .5)$ and the latter is slightly larger than the variance of a uniform distribution on $(-.5, .5)$. The latter seems a more realistic value but we wanted to assess the effect of measurement error on estimated variance components.

The E-M algorithm based on the filtering and smoothing procedures was iterated until the increase in the log likelihood (3.4) was less than .1; the maximum value at convergence was 8017.0. Maximum likelihood estimates of \sum_t and \sum_{e_t} , as in (2.4), are shown in Table 1 for the case where $\sigma_\delta^2 = .10$. Estimates of μ_t are also shown. The units for all variances and covariances are microns². Parametric bootstrap methods (Efron and Tibshirani (1993)) were used to generate standard errors and confidence limits for variance components. Standard errors for estimates of variance tended to be about 10-20% of the size of the estimate. The entire procedure, including 1000 bootstrap replications, used under 7 minutes of CPU time on a DEC OSF/1 V3.2 system when programmed in C++. The estimates obtained when $\sigma_\delta^2 = .04167$ was used were a little different, but the qualitative picture was similar to that in Table 1. The main feature was that $\hat{\sum}_{e_t}$ tended to be about 10% larger than in Table 1, whereas $\hat{\sum}_t$ was more or less the same.

Table 1. Estimated covariance matrices for piston diameters

Stage (t)	$\hat{\sum}_t^1$				$\hat{\sum}_{e_t}^1$				$\hat{\mu}_t$ (mm)
1	4.67	.64	.21	.05	4.67	.64	.21	.05	88.9568
		1.94	.54	.38		1.94	.54	.38	88.9731
			2.29	.51			2.29	.51	88.9350
				3.56				3.56	88.1633
2	5.10	.56	.27	-.10	1.71	.74	.33	.08	88.9574
		3.09	.55	.16		1.36	.35	.18	88.9735
			5.94	.37			3.20	.41	88.9348
				3.24				1.07	88.1616
3	4.69	.42	.08	-.07	2.28	.44	.28	.13	88.9561
		1.43	.44	.01		.81	.44	.37	88.9724
			3.42	.17			2.20	.31	88.9339
				4.64				1.65	88.1601
4	5.91	.64	.18	-.23	2.20	.67	.51	.30	88.9564
		2.13	.46	-.09		1.25	.60	.49	88.9727
			4.79	.22			2.25	.55	88.9346
				4.20				1.92	88.1608

¹The off-diagonal elements are the correlations; the diagonal elements are the variances.

Table 1 suggests that roughly 30-60% of the variation in diameters at each stage is added at that stage and the rest is transmitted from the preceding stage. By using (2.7) recursively we can express $\hat{\sum}_4$ as a sum of four components, one representing the variation at each stage. This indicates that attempts to

reduce variation at the final stage should be directed at stages 3 or 4; little variation is transmitted from stages 1 and 2. We remark that it is also of interest with multivariate measurements to examine their correlation structure. Table 1 indicates a moderate degree of correlation for adjacent diameters in both the total variance and in the variance added at each stage. The examination of principal components or other linear functions of measurement variables is also of general interest but we will not pursue this here.

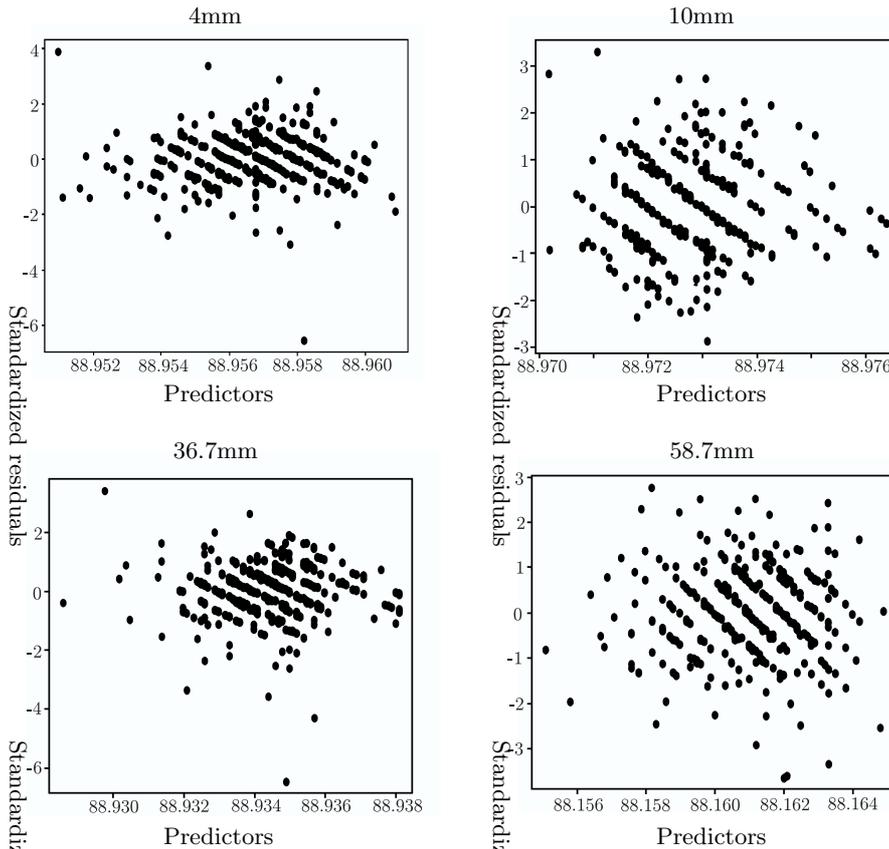


Figure 1. Plots of residuals against the predictors $\hat{y}_{it|t-1}$.

The model (2.1), (2.2) and (2.6) can be checked informally by examining residuals

$$r_{it} = y_{it} - \hat{y}_{it|t-1}$$

or standardized versions of the same. Standardized residuals should look roughly like $N(0, 1)$ variables. Figure 1 shows plots of standardized residuals versus predictors $\hat{y}_{it|t-1}$ across all stage measurements ($i = 1, \dots, 96$; $t = 1, \dots, 4$) for $c = 1, 2, 3, 4$ (corresponding to 4, 10, 36.7 and 58.7 mm). The banded appearance

in each plot is due to the fact that for each diameter there are only 10-15 distinct values of y_{it} , and that the estimated variance for r_{it} does not depend on i and varies slightly with t . Figure 2 shows a normal probability plot of standardized residuals. These are reasonably linear, though a single extreme observation is noted at each of 4 mm and 36.7 mm. More exhaustive checks not shown here likewise do not indicate substantial departures from the working model.

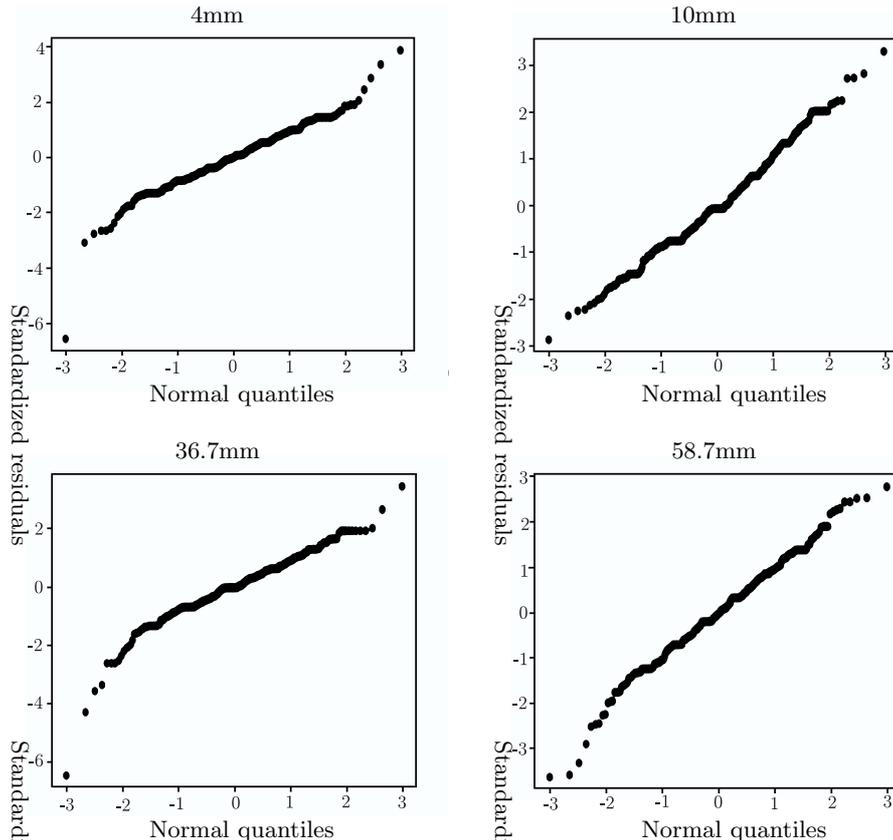


Figure 2. Normal probability plot of standardized residuals.

5. Concluding Remarks

The methods in this paper depend on the approximate validity of a normal AR(1) model for the true measurements. This assumption should be realistic in many contexts, but it would be of interest to consider the implications of model departures. One topic which is readily assessed is the effect of ignoring measurement error. If the model (2.1), (2.2) is assumed correct but there is in fact measurement error as expressed by (2.6), then the maximum likelihood estimates \hat{B}_t derived under (2.1), (2.2) alone converge in probability in large samples not

to B_t but to

$$B_t^* = B_t \sum_{t-1} (\sum_{t-1} + \sum_{\delta_{t-1}})^{-1}.$$

This underestimation of regression parameters is well known when measurement error in covariates is ignored (e.g. Fuller (1987)). A consequence of this in the present circumstances is that the variation transmitted to each stage is underestimated and the variation added is overestimated. This has serious consequences when there are several stages in the process. Agrawal et al. (1996) give a detailed discussion of measurement error for the univariate ($C = 1$) case. They have shown for the case with measurement error but no missing data that the use of simple estimates (3.17) combined with bootstrap confidence intervals provide good procedures. Extension of these methods to the multivariate case is worth considering.

In practical situations one must decide which measurements to consider. This choice can affect whether or not an AR(1) model is satisfactory. For example, if we include a pair of measurements but omit a third which is highly correlated with the other two, we may find an AR(1) model for the two measurements is inadequate.

Further work on ways to interpret multivariate analyses of variation in special contexts is desirable. In particular, one would hope to expose significant relationships among variables and to relate them to the geometry of the units being manufactured. With the piston data there do not appear to be important systematic effects but one could imagine situations in which, for example, the deviations in diameters at opposite ends of a cylinder were negatively correlated after certain stages. The present paper has developed efficient procedures for model fitting and assessment which should make it feasible to undertake further studies with relative ease.

Finally, the methods here deal with processes in which the same variables are measured on parts at each stage. However, as mentioned by Lawless et al. (1996), the general ideas of variation transmission also apply to studies of the effect of upstream process variables on downstream measurements. This area requires further development in practical situations.

Acknowledgements

The authors thank Jeff Robinson for introducing them to variation transmission problems, and Bovas Abraham, Rekha Agrawal, Mike Hamada, Jock MacKay, Diane Gibbons, Jeff Robinson and Mike Wincek for helpful comments on various aspects of this work. Research was partially supported by a fellowship to the first author from The Croucher Foundation and by grants to the second author from the Natural Sciences and Engineering Research Council of Canada and General Motors Canada.

Appendix

Derivation of Filtering and Smoothing Formulas

The filtering formulas (3.6)-(3.9) follow from straightforward conditional mean and variance calculations.

For example,

$$\begin{aligned} z_{it|t-1} &= E\{E[z_{it}|y_{is}, z_{is}, s = 1, \dots, t-1]\} \\ &= E\{A_t + B_t z_{i,t-1}|y_{is}, s = 1, \dots, t-1\} \\ &= A_t + B_t z_{i,t-1|t-1} \end{aligned}$$

$$\begin{aligned} \sum_{iz}(t|t-1) &= E\left\{\sum_{e_t} |y_{is}, s = 1, \dots, t-1\right\} \\ &\quad + \text{Var}\{A_t + B_t z_{i,t-1}|y_{is}, s = 1, \dots, t-1\} \\ &= \sum_{e_t} + B_t \sum_{iz}(t-1|t-1)B_t' \end{aligned}$$

Formulas for $z_{it|t}$, $\sum_{iz}(t|t)$ and the smoothing formulas (3.13), (3.14) are a little more complicated, but may be obtained from standard results about multivariate normal variables. In particular, if x , y and z are random vectors with

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \sim \text{Normal} \left[\begin{pmatrix} \mu_x \\ \mu_y \\ \mu_z \end{pmatrix}, \begin{pmatrix} \sum_{xx} & \sum_{xy} & \sum_{xz} \\ \sum_{yx} & \sum_{yy} & \sum_{yz} \\ \sum_{zx} & \sum_{zy} & \sum_{zz} \end{pmatrix} \right],$$

then

$$\begin{aligned} E(x|z) &= \mu_x + \sum_{xz} \sum_z^{-1}(z - \mu_z) = \mu_{x|z} \\ \text{Var}(x|z) &= \sum_x - \sum_{xz} \sum_z^{-1} \sum_{zx} = \sum_{x|z} \\ \text{Cov}(x, y|z) &= \sum_{xy} - \sum_{xz} \sum_z^{-1} \sum_{zy} = \sum_{xy|z} \end{aligned}$$

and so also, for example,

$$E(x|y, z) = \mu_{x|z} + \sum_{xy|z} \sum_{y|z}^{-1}(y - \mu_{y|z}).$$

Then, for example, letting $y_{i,t-1}^* = (y_{i1}, \dots, y_{i,t-1})'$, we have

$$\begin{aligned} z_{it|t} &= E(z_{it}|y_{it}, y_{i,t-1}^*) \\ &= E(z_{it}|y_{i,t-1}^*) + \text{Cov}(z_{it}, y_{it}|y_{i,t-1}^*) \text{Var}(y_{it}|y_{i,t-1}^*)^{-1}(y_{it} - y_{it|t-1}) \\ &= z_{it|t-1} + \sum_{iz}(t|t-1)H_{it}' \sum_{iy}(t|t-1)^{-1}(y_{it} - y_{it|t-1}). \end{aligned}$$

These formulas are standard in state space models; see for example Harvey and McKenzie (1984) or Koopman and Shepherd (1992).

References

- Agrawal, R., Lawless, J. F. and MacKay, R. J. (1996). Analysis of variation transmission for manufacturing processes in the presence of measurement error. Submitted for publication.
- Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Roy. Statist. Soc. Ser. B* **39**, 1-38.
- Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, London.
- Fuller, W. A. (1987). *Measurement Error Models*. John Wiley, New York.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series, and the Kalman Filter*. Cambridge University Press, Cambridge.
- Harvey, A. C. and McKenzie, C. R. (1984). Missing observations in dynamic econometric models: a partial synthesis. In *Time Series Analysis of Irregularly Observed Data* (Editor by E. Parzen), 108-133. Springer-Verlag, New York.
- Koopman, S. J. and Shepherd, N. (1992). Exact score for time series models in state space form. *Biometrika* **79**, 823-826.
- Lawless, J. F., MacKay, R. J. and Robinson, J. A. (1996). Analysis of variation transmission in manufacturing processes. Submitted for publication.
- Little, R. J. A. and Rubin, D. B. (1987). *Statistical Analysis with Missing Data*. John Wiley, New York.
- Mardia, K. V., Kent, J. T. and Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press.
- Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *Computer Journal* **7**, 308-313.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. (1986). *Numerical Recipes*. Cambridge University Press, Cambridge.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika* **63**, 581-592.
- Shumway, R. H. (1988). *Applied Statistical Time Series Analysis*. Englewood Cliffs, N.J.: Prentice-Hall.

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

E-mail: aytfung@delta.ntu.ac.sg

E-mail: jlawless@setosa.uwaterloo.ca

(Received May 1996; accepted March 1997)