

Center for Quality and Applied Statistics
Kate Gleason College of Engineering
Rochester Institute of Technology

Local-Functional Analysis of Variance for Process Models, Mixture Models, and Deterministic Systems

Joseph G. Voelkel
Center for Quality and Applied Statistics
Rochester Institute of Technology
Joseph.voelkel@rit.edu

Technical Report 2006–1

July 19, 2006



Local-Functional Analysis of Variance for Process Models, Mixture Models, and Deterministic Systems

by

Joseph G. Voelkel

Abstract

Consider the situation in which an experiment is run to study the effects of k factors, at general settings x_1, x_2, \dots, x_k , on a response Y . The factor levels can essentially be fixed and measured at pre-set levels in the experiment, but may exhibit variation from their intended levels in the future. The $\{x_i\}$ include (a) control factors that may be fixed in the future, (b) control factors that may vary within a certain tolerance, and (c) noise factors that may vary to certain hypothetical extents. We consider the case in which the $\{x_i\}$ are continuous factors from either a process experiment or a mixture experiment that can be modeled by a quadratic function, and first examine the total variance of Y induced by the variance of the $\{x_i\}$ settings in the future. We find exact formulas for the total variance of Y in the process case, and compare this to the common first-order propagation-of-error approximation. The method can also be used to find total uncertainty for deterministic equations so long as they can be locally approximated with a quadratic function. We also find “near-exact” formulas for the mixture case, for which we provide formulas under three different mixing scenarios.

We then decompose the variance into additive components due to each of the k factors and into a non-additive portion, which in turn can in turn be decomposed into variance components induced by pairs of factors. We refer to our general approach as local-functional analysis of variance or LANOVA because the decomposition is analogous to components of variance in traditional ANOVA modeling. However, the LANOVA decomposition is a function of the settings’ means (or targets) and variances. We illustrate LANOVA with two response surface models, a mechanistic model, and a mixture model.

Keywords: Design of experiments, tolerance design, parameter design, propagation of error, robust design, variation transmission analysis.

1 Introduction

Experiments have traditionally been designed to find how factors affect the mean of one or more responses. During the last twenty-five years, much attention has been

given to models that incorporate the variance of the response as well. This work was popularized by the robust-design ideas of Taguchi (1987), followed by authors who formulated many alternatives to his work. See, for example, Box and Meyer (1986), León, Shoemaker, and Kacker (1987), Box (1988), Nair and Pregibon (1988), Welch, Yu, Kang, and Sacks (1990), Shoemaker, Tsui, and Wu (1991), and Myers, Khuri, and Vining (1992). However, the notion of studying and reducing the variability of factors' effects may have originated in the ruggedness studies's approach used in the National Bureau of Standards, e.g. Youden (1961), Wernimont (1977), and Paule, R. C. (1988), whose latest documentation is in ASTM E 1169-02 (2002).

Taguchi's formulation called for two sets of experiments. In the first, parameter design, crossed arrays are used with control factors in n treatment combinations (tc's) and, for each such tc, noise factors in the same m combinations. Each control-factor tc produces measures of both location and spread (across noise-factor combinations) and these are used to try to find good settings of control factors. In the second, tolerance design, control factors also play the role of noise factors. Here, the attempt is to find low-cost tolerances.

The typical scenario espoused by Taguchi assumes that there are known costs associated with variation; that there are known costs associated with reduction of uncertainty of control factors; and that the variation associated with noise factors cannot be reduced (equivalent to claiming an infinite cost for such reduction). By contrast, we view control factors that can be set exactly, control factors that can be set within a tolerance, and noise factors that have relatively large variation are simply part of a continuum, each summarized by a measure of center and a measure of uncertainty (and, less frequently, by a measure of association with one or more other factors). We do not explicitly incorporate costs in our analyses for reasons that we note below.

The crossed-array design itself is often criticized, but a number of authors have assumed such a design has been run and have examined better ways to analyze it than Taguchi's method. Vining and Myers (1990) modeled the location and spread response separately, using a quadratic model for each if the design is rich enough to allow it. The spread was modeled directly and all the d.f. for (1) each noise factor, as well as (2) all the noise factors themselves, are grouped together prior to the analysis into an s containing $m - 1$ d.f. for each of the n control-factor runs. We call this approach a *variance-modeled doubly-grouped-d.f. analysis of spread*. Second, say instead that four noise factors were run, each at 3 levels in a 3^{4-1} design. If an s containing 2 d.f. was used to summarize the main effects of each noise factor for each of the control-factor runs, this would be a *variance-modeled singly-grouped-d.f. analysis of spread*. Third, consider Lucas (1994), who "unwound" the crossed array and modeled the mean function, including single d.f. components of the noise factors, and then used a first-order propagation-of-error formula to estimate the variance due to the noise factors. This is a *mean-modeled, first-order, analysis of spread*. See also Myers, Khuri, and Vining (1992). Mathematically, the first method contains the least

amount of information and the third method contains the most. The third method is also the most natural to use when one wishes to consider the effect of trying to control the spread of the noise factors, a situation discussed by Lucas as well as others, e.g. Pignatello and Ramberg in a discussion of León, Shoemaker, and Kacker (1987). Plante (2002) considers a *mean-modeled, second-order, analysis of spread*, with an emphasis on multiple criteria and loss functions.

We extend the mean-modeled, first- and second-order, analysis of spread in two ways, by using a *mean-modeled, exact, decomposed analysis of spread* on linear and quadratic models. The analysis is exact under conditions explained below. The decomposition is performed on the variance of Y analogous to how components of variance of Y are obtained in traditional ANOVA. However, these components of variance are now functions of both the mean (or target, say, for control factors) settings and their associated variance structure. This same idea can be used for deterministic functions that can locally be approximated by a linear or quadratic function. We also extend the idea from process models to mixture models. We refer to this method as *local-functional analysis of variance*, or *LANOVA*. (Note that “functional analysis of variance” is used to describe a global decomposition of deterministic functions in the computer-experimentation literature. See, e.g. Sacks, et al. (1989).)

The ideas we propose are fairly general. While they lend themselves to formal objective functions, we believe that they will be best employed if used to promote inductive learning. For example, before an experiment was run, a noise factor had been declared too expensive to control. However, an analysis showed that this factor had a tremendous impact on results and was not able to be reduced by any of the control factors considered. This analysis led to a revised decision to control the noise factor. Such an insight would likely not have been made if a only formal objective-function approach had been used.

This article assumes information is available to formulate a model. It is not concerned with the important question of how to design experiments to do so.

2 Process-Model Error Structure of the Variables

Let μ_i be a target value for the i^{th} variable, *in the actual scale of the settings*. (For a control factor, such a value is usually determined only after an experiment has been run. For a noise factor, this value is typically set at the mean of the noise factors level and, if so, is known before the experiment is run.) Let x_i be a future realized value, and let $\varepsilon_i \sim \text{ind}(0, \sigma_i^2)$. Later we consider more specific distributions for the $\{\varepsilon_i\}$. The standard structure is

$$x_i = \mu_i + \varepsilon_i.$$

As is customary, we assume that the $\{\sigma_i^2\}$ values are available. Part of the analysis may be to see which, if any, of these values should be reduced (or increased) to produce “good” product at “low” cost. The $\{x_i\}$ are intended to include both control

and noise factors. The $\{\varepsilon_i\}$ are not intended to include measurement error—the $\{x_i\}$ are to be regarded as the actual future values.

There may be occasions where the error structure is more complex. For example, we may set temperature and time, with associated variation, but the model we wish to perturb uses heat and time as factors, where heat is a known function of temperature and time. If this function is linear in temperature and time or, more likely, can be locally approximated as such, this induces a correlation in our error structure. Rather than address this issue here, we simply note that such correlations can be incorporated and we will do so explicitly when we discuss mixture-model error structures. In any event, we write the variance-covariance matrix of $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ as $\mathbf{V}_\mathbf{x}$.

Most models use coded values $\{z_i\}$ instead of the actual values $\{x_i\}$, and we can write a future coded value as

$$z_i = \frac{x_i - M_i}{(U_i - L_i)/2}.$$

Most often, U_i is some upper value used in the experiment, L_i is some lower value, and $M_i = (U_i + L_i)/2$. We write this as $\mathbf{z} = \mathbf{d} + \mathbf{D}\mathbf{x}$, where $d_i = -2M_i/(U_i - L_i)$ and \mathbf{D} is diagonal, with $D_{i,i} = 2/(U_i - L_i)$. In particular,

$$\begin{aligned} \mathbf{E}[\mathbf{z}] &= \boldsymbol{\mu}_\mathbf{z} = \mathbf{d} + \mathbf{D}\mathbf{E}[\mathbf{x}] \\ \text{Var}(\mathbf{z}) &= \mathbf{V}_\mathbf{z} = \mathbf{D}\mathbf{V}_\mathbf{x}\mathbf{D}. \end{aligned}$$

3 Error structure of the response

Let Y^* be the actual response, and let Y be the response *on the analyzed scale*, for example $Y = \ln(Y^*)$. We will assume the model used for Y is either linear or quadratic (full or partial) in \mathbf{z} . In either event we can write the model as

$$Y = \beta_0 + \boldsymbol{\beta}'\mathbf{z} + \mathbf{z}'\mathbf{B}\mathbf{z} + \varepsilon_Y,$$

where β_0 is the intercept, $\boldsymbol{\beta}$ is the vector of the linear coefficients from the model, $\mathbf{B} = (\mathbf{B}_{up} + \mathbf{B}'_{up})/2$ is the symmetric version of \mathbf{B}_{up} , the $m \times m$ upper triangular matrix of the model's second-order coefficients, and ε_Y is a random variable, independent of \mathbf{z} , with mean 0 and variance $\sigma_{Y,\varepsilon}^2$, the unexplained variance of Y if the $\{x_i\}$ were held constant. The coefficients β_0 , those in $\boldsymbol{\beta}$ and \mathbf{B}_{up} , as well as $\sigma_{Y,\varepsilon}^2$, normally need to be estimated for the model and this is typically done through a designed experiment. We also assume that the settings for the i^{th} factor can either be set or recorded exactly in this experiment or that their range is large relative to the joint effects of σ_i and associated coefficients (e.g. Mandel (1984)) so that there is no “errors in measurement” problem. However, in keeping with the usual approach, the uncertainty of these estimates will not be taken into account in the examples that follow—we are concerned here only with point estimates of the variability.

So, with parameters as fixed, but \mathbf{z} and ε_Y as random, we obtain

$$E(Y) = \beta_0 + \boldsymbol{\beta}'\boldsymbol{\mu}_z + \text{tr}(\mathbf{B}\mathbf{V}_z) + \boldsymbol{\mu}_z'\mathbf{B}\boldsymbol{\mu}_z \quad (1)$$

$$\text{Var}(Y) = \text{Var}(\boldsymbol{\beta}'\mathbf{z}) + \text{Var}(\mathbf{z}'\mathbf{B}\mathbf{z}) + 2\text{Cov}(\boldsymbol{\beta}'\mathbf{z}, \mathbf{z}'\mathbf{B}\mathbf{z}) + \sigma_{Y,\varepsilon}^2. \quad (2)$$

where (1) is obtained by $E(\mathbf{z}'\mathbf{B}\mathbf{z}) = E[(\mathbf{z} - \boldsymbol{\mu}_z)'\mathbf{B}(\mathbf{z} - \boldsymbol{\mu}_z)] + \boldsymbol{\mu}_z'\mathbf{B}\boldsymbol{\mu}_z$. First consider the linear case ($\mathbf{B} = \mathbf{0}$). We have

$$E(Y) = \boldsymbol{\beta}'\boldsymbol{\mu}_z \quad (3)$$

$$\text{Var}(Y) = \boldsymbol{\beta}'\mathbf{V}_z\boldsymbol{\beta} + \sigma_{Y,\varepsilon}^2. \quad (4)$$

So, even with the uncertainty in \mathbf{z} , Y is unbiased. Also, $\text{Var}(Y)$ is obtained without any additional assumptions.

Now consider the quadratic case. By (1), Y is “biased” by $\text{tr}(\mathbf{B}\mathbf{V}_z)$, an amount that may be negative, positive or zero, depending on both \mathbf{B} and \mathbf{V}_z , but not $\boldsymbol{\mu}_z$. The variance structure of the higher-order terms involving \mathbf{z} cannot be found directly from \mathbf{V}_z . So, we make the reasonable assumption that ε is multivariate normal (but we modify this slightly for the next section). Then, e.g. Searle (1971, pp. 55–57).
Check Reference

$$\begin{aligned} \text{Var}(Y) = & \boldsymbol{\beta}'\mathbf{V}_z\boldsymbol{\beta} + 2\text{tr}[(\mathbf{B}\mathbf{V}_z)^2] \\ & + 4(\boldsymbol{\mu}_z'\mathbf{B} + \boldsymbol{\beta}')\mathbf{V}_z\mathbf{B}\boldsymbol{\mu}_z + \sigma_{Y,\varepsilon}^2 \end{aligned} \quad (5)$$

Finally, we need to translate this back to the original response Y^* . If $Y = Y^*$, we are done. If Y is transformed according to, say, the two-parameter Box-Cox (1964) transformation

$$Y = \begin{cases} (Y^* + k)^\lambda, & \lambda \neq 0 \\ \ln(Y^* + k), & \lambda = 0 \end{cases},$$

then, to a first-order expansion,

$$\text{Var}(Y^*) = \begin{cases} \text{Var}(Y) \lambda^{-2} [E(Y)]^{-2+2/\lambda}, & \lambda \neq 0 \\ \text{Var}(Y) \exp[2E(Y)], & \lambda = 0 \end{cases}.$$

Other transformations are handled in a similar manner. Thus, the variance in \mathbf{x} gets transmitted to the variance in Y^* through the path $\mathbf{x} \rightarrow \mathbf{z} \rightarrow Y \rightarrow Y^*$, whose key section is $\mathbf{z} \rightarrow Y$. In the next section, decomposition of this variance is explained.

4 Process-Model Quadratic Example

To gain insight into (2), we express the variance of Y in terms of that of \mathbf{z} directly for the two-factor full-quadratic model:

$$Y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_{11} z_1^2 + \beta_{22} z_2^2 + \beta_{12} z_1 z_2 + \varepsilon_Y.$$

Let z_i be independent random variables with means ν_i (usually the targets, for control factors) and variances τ_i^2 , where (ν_i, τ_i^2) is used in this section to simplify notation. Assume for simplicity that both z_i are sampled from symmetric distributions with proper kurtoses denoted by κ_i , i.e., $\kappa_i = 3$ for the normal distribution. In addition to the variances in (4), these assumptions generate the following non-zero variance-covariance structures: $\text{var}(z_i^2) = (\kappa_i - 1)\tau_i^4 + 4\nu_i^2\tau_i^2$, $\text{var}(z_1 z_2) = \tau_1^2\tau_2^2 + \nu_1^2\tau_2^2 + \nu_2^2\tau_1^2$, $\text{cov}(z_i, z_i^2) = 2\nu_i\tau_i^2$, $\text{cov}(z_i, z_i z_j) = \nu_j\tau_i^2$, $\text{cov}(z_i^2, z_i z_j) = 2\nu_j\nu_i\tau_i^2$.

Then, aside from the variance $\sigma_{Y,\varepsilon}^2$ due to ε_Y , we can decompose the total variance $\sigma_{Y,Tot}^2$ of Y as follows. The pure contribution from z_1 is

$$\begin{aligned} \sigma_{Y,1}^2 &= \tau_1^2 (\beta_1^2 + 4\beta_{11}^2\nu_1^2 + \beta_{12}^2\nu_2^2 + 4\beta_1\beta_{11}\nu_1 + 2\beta_1\beta_{12}\nu_2 + 4\beta_{11}\beta_{12}\nu_1\nu_2) \quad (6) \\ &\quad + \tau_1^4\beta_{11}^2(\kappa_1 - 1) \\ &= \tau_1^2 (\beta_1 + 2\beta_{11}\nu_1 + \beta_{12}\nu_2)^2 + \tau_1^4\beta_{11}^2(\kappa_1 - 1), \end{aligned}$$

the pure contribution from z_2 , $\sigma_{Y,2}^2$, follows from symmetry, and the non-additive contribution, from the joint effects of z_1 and z_2 , is $\sigma_{Y,N}^2 = \tau_1^2\tau_2^2\beta_{12}^2$. Note that (6) is a no-intercept quadratic function of τ_1^2 , may be a function of both ν_1 and ν_2 , and includes a contribution due to β_{12} . Also note that the non-additive contribution is constant in both ν_1 and ν_2 .

Software packages that use first-order propagation of error formulas do not include the sum $\tau_1^4\beta_{11}^2(\kappa_1 - 1) + \tau_2^4\beta_{22}^2(\kappa_2 - 1) + \tau_1^2\tau_2^2\beta_{12}^2$ in their formula, so for quadratic models they underestimate the total variance. This is especially noticeable at the stationary point, when the first-order solution predicts 0 variance due to the uncertainty in \mathbf{z} . Similarly, if the stationary point is a maximum, say, then either intuitive considerations or use of (1) applied to the canonical problem show that $E(Y)$ is less than $E(Y|\mathbf{z})$ at the stationery point.

For the general case, one can obtain $\sigma_{Y,i}^2$ by setting $\sigma_{Y,\varepsilon}^2 = 0$ and $\sigma_k^2 = 0$ for $k \neq i$ and applying (5) for each i . To find the total non-additive contribution, apply (5) with only $\sigma_{Y,\varepsilon}^2 = 0$ and then subtract out $\sum_i \sigma_{Y,i}^2$. If this contribution is small it would normally not need to be decomposed. If it is large, it can be decomposed into i, j terms by using $\sigma_{Y,\varepsilon}^2 = 0$ and $\sigma_k^2 = 0$ for $k \neq i, k \neq j$, for all $i < j$, and subtracting out the pure contributions from i and j .

Taylor (1991), whose work on variation transmission analysis we discovered after formulating our results, is in the spirit of our work. Taylor emphasized standard deviations instead of variances, considered the special case when $\mathbf{V}_{\mathbf{z}}$ is a diagonal

matrix, did not consider mixtures, and used a second-order Taylor series expansion—this does not explicitly account for kurtoses, but effectively assumes kurtoses of 2.

5 Mixture-Model Error Structures

The mixture-model error structure is more complex than the process-model error structure for two reasons:

1. An inherent correlation exists in the error structure because the levels of the variables, or components, must sum to a constant (usually taken to be 1).
2. The nature of how components are added may induce an additional correlation structure.

We first consider a general situation and then provide three examples. To make the situation concrete, assume we are adding components by their weights. Let μ_i be the target weight of the i^{th} component and let u_i^* be a future realized weight. We will start with the general model

$$u_i^* = \mu_i + \eta_i, \quad i = 1, 2, \dots, m; \quad \boldsymbol{\eta} \sim (\mathbf{0}, \mathbf{V}),$$

where \mathbf{V} is a general variance-covariance matrix. The intended total weight is $\sum_i \mu_i = \mu_{Tot}$, and the realized total weight is $\sum_i u_i^* = u_{Tot}^*$.

We consider three error-structure scenarios. For each, $\varepsilon_i \sim \text{ind}(0, \sigma_i^2)$.

Independent additions. Here, the weights of each component are measured separately on one or more scales. A reasonable model is

$$u_i^* = \mu_i + \varepsilon_i,$$

so $\eta_i = \varepsilon_i$ and $\mathbf{V} = D(\sigma_i^2)$, a diagonal matrix with $D_{ii} = \sigma_i^2$. (If a wide variety of target weights is to be considered for each component, it may be reasonable to use $\sigma_i = k_i \mu_i$, e.g., instead.)

Cumulative additions. An example of this occurs when components are added sequentially into a beaker that has been tared. If the taring error is ε_0 , then the first, second, \dots , m^{th} weights realized can be modeled as

$$\begin{aligned} u_1^* &= \mu_1 + \varepsilon_1 - \varepsilon_0 \\ u_1^* + u_2^* &= \mu_1 + \mu_2 + \varepsilon_2 \\ &\vdots \\ u_1^* + u_2^* + \dots + u_m^* &= \mu_1 + \mu_2 + \dots + \mu_m + \varepsilon_m, \end{aligned}$$

where the weighing errors are the $\{\varepsilon_i\}$ and $\sigma_i^2 = \sigma^2$, $i = 0, \dots, m$ is often reasonable. Rearranging terms produces

$$\begin{aligned} u_1^* &= \mu_1 + \varepsilon_1 - \varepsilon_0 \\ u_2^* &= \mu_2 + \varepsilon_2 - \varepsilon_1 \\ &\vdots \\ u_m^* &= \mu_m + \varepsilon_m - \varepsilon_{m-1}, \end{aligned}$$

so $\eta_i = \varepsilon_i - \varepsilon_{i-1}$. As a result, \mathbf{V} is a banded matrix with $V_{i,i} = \sigma_i^2 + \sigma_{i-1}^2$, $V_{i,j} = -\sigma_{\min(i,j)}^2$ for $|i - j| = 1$ and is 0 otherwise. If $m = 4$,

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 + \sigma_0^2 & -\sigma_1^2 & 0 & 0 \\ -\sigma_1^2 & \sigma_2^2 + \sigma_1^2 & -\sigma_2^2 & 0 \\ 0 & -\sigma_2^2 & \sigma_3^2 + \sigma_2^2 & -\sigma_3^2 \\ 0 & 0 & -\sigma_3^2 & \sigma_4^2 + \sigma_3^2 \end{bmatrix}.$$

All-but-last independent additions. This is the same as the first scenario, except that the last component (often water or a similar filler) is added to achieve a certain (measured) weight. Then

$$\begin{aligned} u_i^* &= \mu_i + \varepsilon_i, \quad i = 1, \dots, m - 1 \\ u_1^* + u_2^* + \dots + u_m^* &= \mu_1 + \mu_2 + \dots + \mu_m + \varepsilon_m \end{aligned}$$

so

$$\begin{aligned} u_i' &= \mu_i + \varepsilon_i, \quad i = 1, \dots, m - 1 \\ u_m' &= \mu_m - \sum_{i < m} \varepsilon_i + \varepsilon_m, \end{aligned}$$

and this leads to $V_{i,i} = \sigma_i^2$ for $i < m$, $V_{m,m} = \sum_{i=1}^m \sigma_i^2$, $V_{i,m} = V_{m,i} = -\sigma_i^2$ for $i < m$, and otherwise $V_{i,j} = 0$. If $m = 4$,

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & -\sigma_1^2 \\ 0 & \sigma_2^2 & 0 & -\sigma_2^2 \\ 0 & 0 & \sigma_3^2 & -\sigma_3^2 \\ -\sigma_1^2 & -\sigma_2^2 & -\sigma_3^2 & \sum_{i=1}^4 \sigma_i^2 \end{bmatrix}$$

Other error structures may exist and \mathbf{V} can be found in a similar manner.

The model used in mixtures is such that the components must add up to a fixed amount, say a_{Tot} . Most often $a_{Tot} = 1$, but the user might only consider a model in which a fixed fraction will be used, or may prefer to express the components on a percent scale. We introduce intermediate variables $\{u_i\}$ here that are simply rescaled

versions of the $\{u_i^*\}$. Let $r_{\mu a} = a_{Tot}/\mu_{Tot}$

$$\begin{aligned} a_i &= \mu_i r_{\mu a} \\ u_i &= u_i^* \frac{a_i}{\mu_i} = u_i^* r_{\mu a} \\ u_{Tot} &= \sum_i u_i \end{aligned}$$

The variance-covariance matrix of \mathbf{u} is simply $r_{\mu a}^2 \mathbf{V}$. Also note that the $\{u_i\}$ are not restricted to add to a fixed amount, unless this is imposed on them by the error structure. Such an imposition is equivalent to $\mathbf{1}'\mathbf{u}$ being constant, or $0 = \text{Var}(\mathbf{1}'\mathbf{u}) = \mathbf{1}'\mathbf{V}\mathbf{1}$, or $\sum_{i,j} V_{i,j} = 0$. This occurs under our scenarios in the following ways:

1. *Independent Addition.* All $\sigma_i^2 = 0$, a trivial and uninteresting condition.
2. *Cumulative additions.* This holds if there is no taring error, and if the last component is added without error, i.e., $\sigma_0^2 = \sigma_m^2 = 0$.
3. *All-but-last independent additions.* This holds if the last component is added without error, i.e., $\sigma_m^2 = 0$.

6 Mixture and pseudo components and their error structures

The mixture components $\{x_i\}$ are simply the $\{u_i\}$, but rescaled stochastically so their sum is *exactly* a_{Tot} . That is,

$$x_i = a_{Tot} \frac{u_i^*}{u_{tot}^*} = a_{Tot} \frac{u_i}{u_{tot}},$$

forcing $\sum_i x_i = a_{Tot}$.

We now approximate the variance-covariance matrix of \mathbf{x} by using a first-order Taylor series expansion. To do this, let $g_i(\mathbf{u}) = a_{Tot} u_i / u_{tot}$, so

$$\begin{aligned} x_i &= g_i(\mathbf{u}) \\ &\simeq g_i(\mathbf{a}) + \sum_{k=1}^m (u_k - a_k) \left. \frac{\partial g_i(\mathbf{u})}{\partial u_k} \right|_{\mathbf{u}=\mathbf{a}} \\ &= a_i + \sum_{k \neq i} (u_k - a_k) (-a_i / a_{Tot}) + (u_i - a_i) (1 - a_i / a_{Tot}) \end{aligned} \quad (7)$$

To the approximation (7), $E[x_i] = a_i$, its target. Let $f_i = a_i / a_{Tot}$ be the target

fraction. Then, to the same approximation, some algebra produces

$$\begin{aligned} \text{Var}(x_i)/r_{\mu a}^2 &= f_i^2 \sum_{k \neq i} V_{k,k} + V_{i,i} (1 - f_i)^2 \\ &\quad + 2f_i^2 \sum_{k < l; k, l \neq i} V_{k,l} - 2f_i (1 - f_i) \sum_{k \neq i} V_{k,i} \end{aligned}$$

and

$$\begin{aligned} \text{Cov}(x_i, x_j)/r_{\mu a}^2 &= f_i f_j \sum_{k \neq i} \sum_{l \neq j} V_{k,l} \\ &\quad + (1 - f_i)(1 - f_j) V_{i,j} \\ &\quad - f_j (1 - f_i) \sum_{l \neq j} V_{i,l} \\ &\quad - f_i (1 - f_j) \sum_{k \neq i} V_{j,k}. \end{aligned}$$

Assembling these values, we can write $\text{Var}(\mathbf{x}) = \mathbf{V}_x$, the error structure of the $\{x_i\}$. Note that this structure depends on the error structure of the $\{u_i\}$, the fractions of the $\{x_i\}$ in the mixture, and the scale factor $r_{\mu a}$.

The $\{x_i\}$ are the realized actual components. But most statistical mixture techniques transform these to pseudo-components $\{z_i\}$, where

$$z_i = \frac{x_i - L_i}{a_{Tot} - \sum L_i},$$

where L_i is the lower bound for x_i in the particular experiment run. We can write this linear transformation as $\mathbf{z} = \mathbf{d} + \mathbf{D}\mathbf{x}$, where $d_i = -L_i/(a_{Tot} - \sum L_i)$ and \mathbf{D} is diagonal, with $D_{k,k} = 1/(a_{Tot} - \sum L_i) = r_{xz}$. In particular,

$$\begin{aligned} \mathbf{E}[\mathbf{z}] &= \boldsymbol{\mu}_z = \mathbf{d} + \mathbf{D} \mathbf{E}[\mathbf{x}] = \mathbf{d} + \mathbf{r}_{xz} \mathbf{E}[\mathbf{x}] \\ \text{Var}(\mathbf{z}) &= \mathbf{V}_z = \mathbf{D} \mathbf{V}_x \mathbf{D} = r_{xz}^2 \mathbf{V}_x \end{aligned}$$

For the error structure of the response, only minor modifications are needed for the mixture case. For \mathbf{z} , we will use the Scheffé parametrization. For $Y = \beta_0 + \boldsymbol{\beta}'\mathbf{z} + \mathbf{z}'\mathbf{B}\mathbf{z} + \varepsilon_Y$, we drop β_0 . For $\mathbf{B} = (\mathbf{B}_{up} + \mathbf{B}'_{up})/2$ the \mathbf{B}_{up} becomes a strictly upper triangular matrix, so it does not include “pure quadratic” terms.

The path is $\mathbf{u}^* \rightarrow \mathbf{u} \rightarrow \mathbf{x} \rightarrow \mathbf{z} \rightarrow Y \rightarrow Y^*$, and the key sections are $\mathbf{u} \rightarrow \mathbf{x}$ and $\mathbf{z} \rightarrow Y$.

Steiner and Hamada (1997) considered the independent-additions problem. However, they noted that, “due to the interaction of measurement errors, the effect of mixing measurement errors is complex and no simple closed form expression ... can

be obtained.” They suggested a simulation of the mixing measurement error pattern. We have provided such an expression. See Hamada, Martz, and Steiner (2005) for a related problem.

7 Example 1. Process Model to Target

We provide four examples, designed to illustrate different features of the LANOVA approach. Although we expect our approach to be useful for the noise-factor case, none of the examples include them—the experiments we have seen that involve noise factors are not rich enough to support a full (or even interesting partial) quadratic model.

Our first example is based on the “rsm-a” example data in Design-Expert[®]. This a 3-factor CCD, for which the reduced model for the response Conversion (as a percent) was

$$\begin{aligned}\hat{Y} = & 80.04 + 1.03z_1 + 4.04z_2 + 6.20z_3 + \\ & 3.07z_2^2 - 5.07z_3^2 + 11.38z_1z_3 - 3.87z_2z_3.\end{aligned}$$

When un-coded, the three variables had “ ± 1 ” settings whose ranges were 10, 10, and 1 respectively. Suppose we wish to produce conversion to a target of 90% and that the variables could be controlled to $\sigma_i = 0.5, 0.5, 0.1$, respectively, on this scale (or $(0.1, 0.1, 0.2)$ on the z -scale). The desirability-function method (Derringer and Suich, 1980) was used to minimize the tolerance variance while maintaining a target setting. For this and other examples, we will ignore the contribution of the residual variance—this measure from a designed experiment often does not translate well into the future $\sigma_{Y,\varepsilon}^2$ of interest (and for this problem turned out to be large relative to the other sources of variance we wish to consider). Seventy solutions were obtained using Design-Expert, based on slight variations entered in the desirability criteria and also by random starting points. These were reduced to 16 solutions using hierarchical clustering on the z vectors to eliminate solutions that were very close in design space.

To examine a simpler case first, pretend these 16 z solutions were obtained as if the model above did not include the two interaction terms. The LANOVA was performed at each of these conditions, by using (5) and the method explained in Section 4. These results are shown in the top graph in Figure 1, which is sorted by the dominant variance component, due to variable 3. This graph, in conjunction with the bottom graph, reveals that the higher variances due to variable 3 are associated with lower z_3 settings. The opposite holds for variable 2. The variance contribution to variable 1 is uniformly low. Because the model we are assuming has no interactions, each component makes a contribution to the total that is independent of where the the other factors are set. This can be seen in Figure 2, which plots each variance component $\sigma_{Y,i}^2$ in Figure 1 versus its settings.

For the actual problem, with interactions in the model, the solutions obtained are

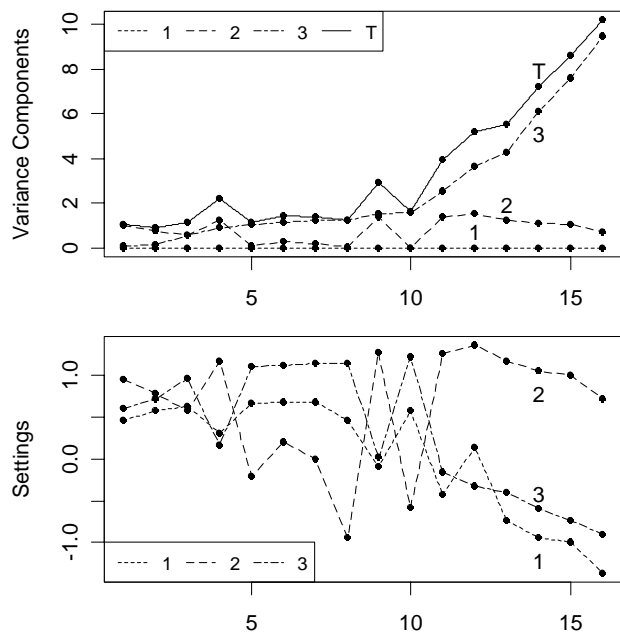


Figure 1: Rsm-a Example, Using 16 z Points but Assuming No Interaction in the Model. LANOVA Plot, Sorted by Component 3, and Plot of Settings.

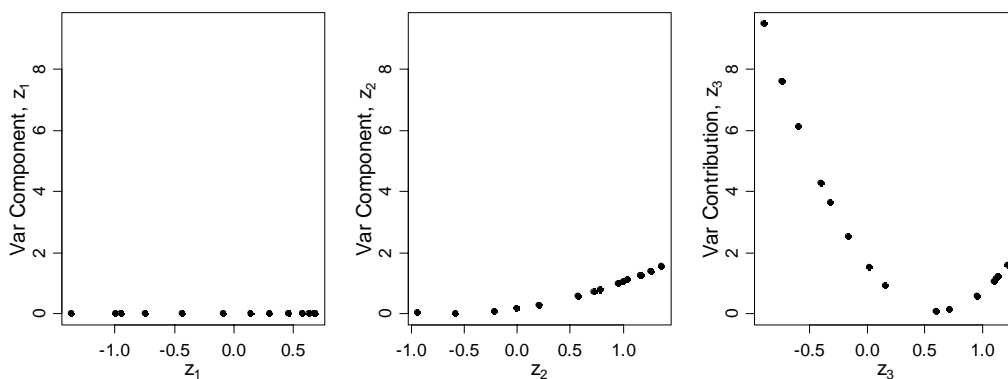


Figure 2: Rsm-a Example, Assuming no Interaction in the Model. Estimates of Variances σ_i^2 .

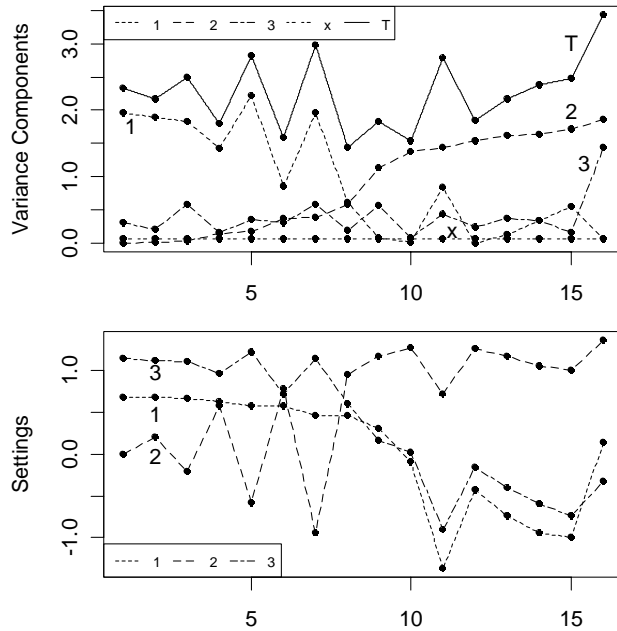


Figure 3: Rsm-a Example, with Interactions in the Model. LANOVA Plot, Sorted by Component 2, and Plot of Settings.

shown in Figure 3, which is again sorted by the dominant variance component, now due to variable 2. The key feature in the LANOVA graph is that there appear to be at least two distinct ways in which equal means and variances for Y may be found. The lower graph indicates these locations in \mathbf{z} space. Figure 4 shows the effects of the interactions in the model.

The variance contributions due to z_1 and z_2 are shown in Figure 5.

This figure suggests a better way to order the settings than by simply sorting by the second variance component—order the settings by the trace suggested in the figure. If we start the indexing at the lowest value of z_1 and, for smoothness, remove the unusual value as seen in Figure 4, we achieve Figure 6, which illustrates the continuous nature of the results. This figure in which an apparently one-dimensional set of solutions was found, may be regarded as an extension of the standard trace plot for process variables. However, in the more general case a two- or higher-dimensional set of solutions may be found, and these could be regarded as an extension of trace plots to higher dimensions.

Next, consider the settings with index values of 5 and 11 in Figure 3, which are summarized in Table 1. An experimenter who needs to reduce the total variance down to, say, 2.00, but who only has the total variance available, would likely select

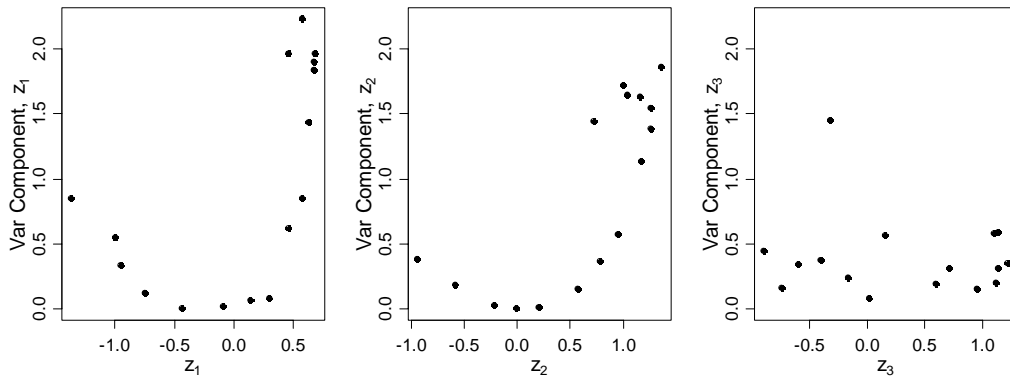


Figure 4: Rsm-a Example, with Interactions in the Model. Estimates of Variances σ_i^2 .

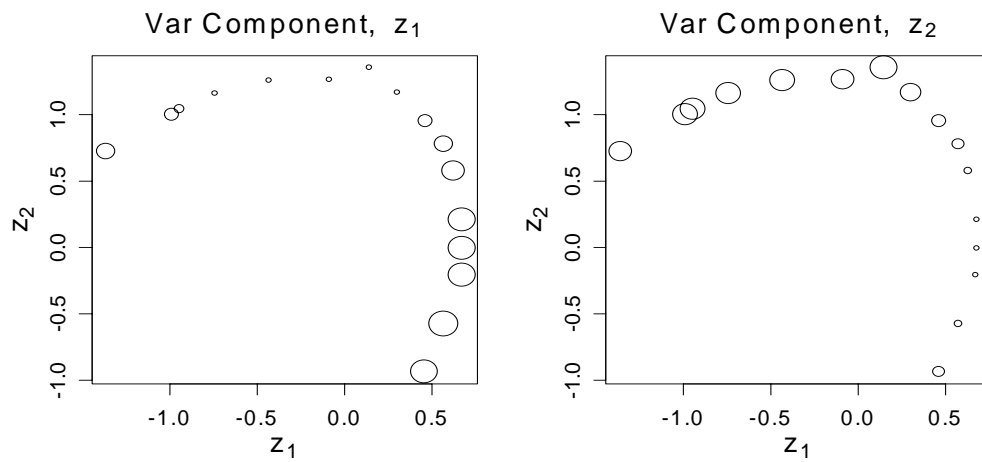


Figure 5: Rsm-a Example, with Interactions in the Model. Areas of Circles Represent Estimated Amount of Variance Contributions due to z_1 (Left Graph) and z_2 (Right Graph).

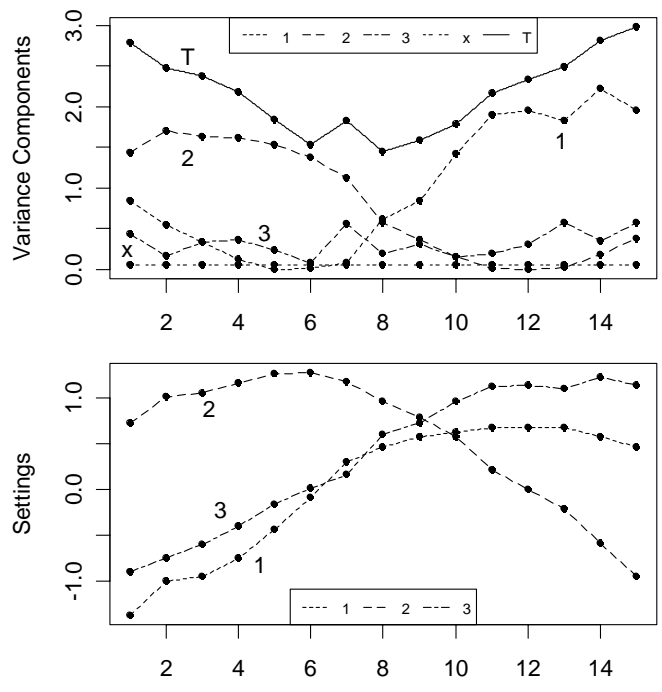


Figure 6: Rsm-a Example, with Interactions in the Model. LANOVA Plot, Sorted by Trace from Figure 5.

	Setting 5	Setting 11
z_1	0.57	-1.36
z_2	-0.58	0.73
z_3	1.22	-0.90
$\sigma_{Y,1}^2/\sigma_{Y,Tot}^2$	79%	30%
$\sigma_{Y,2}^2/\sigma_{Y,Tot}^2$	7%	52%
$\sigma_{Y,3}^2/\sigma_{Y,Tot}^2$	12%	16%
$\sigma_{Y,N}^2/\sigma_{Y,Tot}^2$	2%	2%
$\sigma_{Y,Tot}^2$	2.81	2.78

Table 1: Rsm-a Example, with Interactions. LANOVA Summary of Settings 5 and 11.

the setting with the lower total, setting 11. But whichever setting is selected still does not provide any information about what to do next—the experimenter would need to change the σ_i by different amounts to “see what happens.” By contrast, using LANOVA, it is clear that a reduction to $\hat{\sigma}_{Y,Tot}^2 = 2.00$ can be obtained by reducing the total percent contribution by $1 - 2.00/2.81 = 29\%$. This could be done in setting 5 by reducing $\hat{\sigma}_1^2$ to $50/79 = 63\%$ of its original value, or in setting 11 by either virtually eliminating the variance from variable 1 or by reducing $\hat{\sigma}_2^2$ to $23/52 = 44\%$ of its original value. The LANOVA also makes it clear that for either setting, reduction of the $\hat{\sigma}_3^2$ alone will be useless.

Note that these results hold because the contribution due to the non-additive portion is quite small, a feature that should hold unless a stationary or near-stationary point is chosen on a surface with large interactions (our next example). Also, this is an example where it may be advantageous *not* to consider costs directly in a loss function, especially if this is in an early stages of an investigation. The LANOVA illustrates directly the various options available to the engineer. By contrasts, loss-function costs are often crudely and artificially calculated and can frequently obscure key issues—to pretend they are known can narrow down the investigation at too early a stage and produce a formal solution that is neither useful nor optimal.

8 Example 2. Process Maximum

The second example is based on an experiment discussed in Box (1954) and Box and Draper (1987, p. 368), in which five factors (temperature T_1 , time t_1 as $\log(t_1)$, and a concentration reactant C at the first stage, and temperature T_2 and time t_2 as $\log(t_2)$ at the second stage) were studied to improve the yield of a chemical reaction. The uncertainties are not given, but we use $\sigma_z = 0.1$ for each standardized variable.

The interesting features of this experiment were (1) the optimum yield existed inside the experimental region, (2) the optimum could be well approximated by a two-dimensional ridge system, and (3) the $\log(t_1) \times T_1$, $C \times T_1$ and $C \times T_2$ interactions

	Original	Canonical
$\sigma_{Y,1}^2/\sigma_{Y,Tot}^2$	9%	66%
$\sigma_{Y,2}^2/\sigma_{Y,Tot}^2$	6%	23%
$\sigma_{Y,3}^2/\sigma_{Y,Tot}^2$	22%	10%
$\sigma_{Y,4}^2/\sigma_{Y,Tot}^2$	18%	1%
$\sigma_{Y,5}^2/\sigma_{Y,Tot}^2$	7%	0%
$\sigma_{Y,N}^2/\sigma_{Y,Tot}^2$	38%	0%
$\sigma_{Y,12}^2/\sigma_{Y,Tot}^2$	6%	
$\sigma_{Y,13}^2/\sigma_{Y,Tot}^2$	7%	
$\sigma_{Y,34}^2/\sigma_{Y,Tot}^2$	21%	
$\sigma_{Y,other\ ij}^2/\sigma_{Y,Tot}^2$	4%	
$\sigma_{Y,Tot}^2$	0.00603	0.00603

Table 2: Yield Example. LANOVA, with Original and Canonical Variables at Optimum settings.

were quite large, accounting for about one-third of the total R_{adj}^2 among the effects from an ANOVA.

Consider the uncertainty at the optimum setting (that which led to the maximum estimated yield). Indexing the original variables $T_1, \dots, \log(t_2)$ from 1 to 5 produces the LANOVA shown in the second column of Table 2. Note that the non-additive portion accounts for 38% of the total induced variance, due to the strong influence of the interactions at the optimum setting. These were decomposed as well, as explained in Section 4. It is quite clear that any real reduction in variability must come from factors 3 and 4 (C and T_2). If we eliminate all variability due to C alone, the reduction in induced variance is 50% (22% + 7% + 21%); for D alone, the reduction is 39%; and for both, the reduction is 68%.

By comparison, the canonical variables, listed in order of decreasing modulus of their eigenvalues, highlight the two-dimensional ridge system (small contribution from 4th and 5th principle components) and, of course, no contribution (algebraic zero) due to non-additive effects. All of the contributions for both original and canonical variables would be 0 under first-order propagation-of-variance approximations.

We also note that the “2 degrees of freedom” for the ridge system was used by Box to find settings in which the total reaction time $t_1 + t_2$ could be kept low while still keeping yield at or near its maximum value. This leaves “1 degree of freedom” that could be used to minimize the induced variation and to learn more about the contributions due to individual, or pairs of, contributions.

Feature	Nominal	Systematic Uncertainty B_i	Random Uncertainty σ'_i
C discharge coefficient	0.964	0.0075	0
d Throat diameter (in)	4	0.001	0
D Inlet diameter (in)	6	0.002	0
ρ Water density, 60°F (lbm/ft ³)	62.37	0.004	0.002
h Differential pressure head (in H ₂ O)	100	0.3	0.4

Table 3: Venturi Features (Uncalibrated)

9 Example 3. ASME Test Uncertainty

Section 10.2 of ASME (1998) considers the uncertainty in the flow of water using a 6×4 in. venturi. The objective was to see whether an uncalibrated venturi could be used to satisfy the test objective and, if not, whether the calibration of the venturi would be sufficient.

Each feature of the venturi is subject to either systematic and/or random uncertainty, where the systematic uncertainty, a summary of bias, is in accord with the ISO Guide (1993). Feature information is summarized in Table 3.

The flow rate (lbm/s) is given by

$$m = \frac{0.099702Cd^2\sqrt{\rho h}}{\sqrt{1 - (d/D)^4}}.$$

The flow rate's total uncertainty (a term we explain later) is desired to be no greater than 0.5% of the flow. In the ASME manual, this is done by use of first-order Taylor series expansions, and by keeping the systematic and random components separate until the end of the analysis. No quantification is given of the relative contribution of each component. By comparison, we proceed as follows:

1. Formally, define $\sigma_i^2 = B_i^2/2 + \sigma_i'^2$ as the *variability* associated with the i^{th} feature. This formulation allows the following to be consistent with the ISO standards.
2. Calculate m in the region of interest using a design that is rich enough to fit a quadratic model. We chose to do this by using a 3^5 factorial array. To investigate the linearity assumption we set the levels for each factor at Nominal $\pm 4\sigma$. At the nominal value, this gives $m = 138.39$ lbm/s.
3. Fit the model. The model was highly linear in the range of interest (ANOVA $R_{Adj}^2 = 0.999990$), although the addition of 2^{nd} -order terms technically im-

Feature	% Total	% Systematic	% Random
C discharge coefficient	75.7	75.7	0.0
d Throat diameter	0.5	0.5	0.0
D Inlet diameter	0.0	0.0	0.0
ρ Water density	0.0	0.0	0.0
h Differential pressure head	23.8	2.9	20.8
Total	100.0	79.2	20.8

Table 4: Venturi Example. LANOVA Percent Contributions of Variability, Uncalibrated Venturi

proved the fit. For illustrative purpose, the full quadratic model was maintained. Note that our method can naturally check the linearity assumption as part of the analysis, whereas the ASME approach assumes it. (However, the ASME method covers situations that we do not consider, such as different amounts of uncertainty above and below the target settings.)

4. Find the variability induced by each of the components on the response as well as the total. The total variability is $\sigma_{Y,Tot}^2 = 0.3674 \text{ (lbm/s)}^2$. This converts to a total uncertainty, calculated (roughly based on 95% limits) as $2\sqrt{0.3674} = 1.21 \text{ lbm/s}$, or a relative total uncertainty of $1.21/138.39 = 0.87\%$. This result (also given in ASME) is larger than the 0.5% objective.
5. Split the induced variability into sub-components due to systematic and random components by using ratios based on step (1.). These are expressed as percent contributions in Table 4.

In the ASME example, it was felt that the systematic component of the discharge coefficient could be reduced from 0.0075 to 0.0012 by calibrating the venturi. Will this be enough to obtain the 0.5% objective? We can calculate this directly from Table 4. We need to eliminate at least $1 - (0.5/0.87)^2 = 67.0\%$ of the variability, so we would need to reduce the systematic component to $0.0075\sqrt{1 - 67.0/75.7} = 0.0025$. Hence, a reduction to 0.0012 will meet the objective. This is a reduction of 97.4% of the original variability, from 75.7% to 1.9%, so the revised total uncertainty would become $0.87\%\sqrt{100 - (75.7 - 1.9)} = 0.45\%$.

This revised total can, of course, be calculated directly as well, by replacing 0.0075 with 0.0012 in Table 3 and proceeding as before—this is approach used in the ASME manual. However, our table can lead to more insight into the problem, especially for more complex cases. For example, it is clear from the table that any further serious reductions in uncertainty must now come from the differential pressure head, which, with the variability reduction used above, would now account for over 90% of the total remaining variability.

Component	Low (%)	High (%)	Target (g)	s.d. (g)
<i>A</i>	5	15	89	2
<i>B</i>	10	20	148.3	6
<i>C</i>	0	10	48.3	2
<i>D</i>	1	4	23.7	2
<i>E</i>	42.5	75.5	605.7	10

Table 5: Mixture Example. Component Information

10 Example 4. Mixture Experiment

Our last example is based on a liquid formulation involving 5 components. The experiment is proprietary, so components will be labeled A–E. The design was a constrained mixture, 30-run, D-optimal (for quadratic) design. The five components were added independently as part of a 1000g mixture, of which the components comprised 915g. To use the independent-additions case here, we assume that the final mixture does contain exactly 91.5% of the full mixture. Low and high settings for the components appear in columns 2 and 3 of Table 5. Two responses were measured, of which we only consider one, viscosity.

Based on requirements for the two response and their corresponding models from the experiment, the target settings for the components were selected and are shown in column 4 of Table 5. The s.d.’s for the components are also shown. (Their relatively large values indicated a desire to create the mixture quickly with low-cost equipment.) In our notation, the fourth column of the table refers to the μ_i ’s and the fifth column to the σ_i ’s. Also, $\mu_{Tot} = 915$, $a_{Tot} = 91.5$. The x_i ’s, the components associated with u_i^* ’s but that are stochastically rescaled to sum to a_{Tot} , have a variance-covariance matrix \mathbf{V}_x that will be summarized here by a matrix whose upper diagonal elements are correlations and whose diagonal elements are standard deviations as shown in (8). Note that if the x_i ’s were uncorrelated, these standard deviations would simply be those from Table 5, divided by 10. The restriction on the sum has made some of these standard deviations on the diagonal larger, and some smaller (especially component *E*). Also, one might anticipate that all correlations would be negative, but that is not the case.

$$\begin{bmatrix} 0.21 & -0.16 & 0.04 & -0.03 & -0.23 \\ & 0.53 & -0.12 & -0.09 & -0.80 \\ & & 0.20 & -0.03 & -0.24 \\ & & & 0.20 & -0.24 \\ & & & & 0.57 \end{bmatrix} \quad (8)$$

An analysis indicates the following model (coded in pseudo-components) is rea-

sonable for viscosity ($R_{Adj}^2 = 98.5\%$):

$$\begin{aligned}\hat{Y} = & 4.1159A + 2.8236B + 2.4507C + 7.9120D + 1.4591E \\ & -2.0950AE - 1.1875BE - 5.4919DE\end{aligned}$$

To find the variance contributions due to all components, we proceed as before. However, to find the variance associated purely with the i^{th} component, it is useful to construct not only \mathbf{V}_z , but also $\mathbf{V}_{z,i}$, the variance-covariance matrix obtained by assuming that only component i has a non-zero σ_i . This provides a simple way to find the contribution that includes only σ_i terms. This example reveals that if only $\sigma_2 = 6$ is non-zero, the following s.d.'s are induced upon the x_i 's: (0.06, 0.50, 0.03, 0.02, 0.40). In particular, the ‘‘anticipated’’ $\sigma_2/10 = 0.6$ is reduced to 0.50, but the precisely measured component 5 has a surprisingly large s.d. In fact, if instead only $\sigma_5 = 10$ is non-zero, the resulting s.d.'s of the x_i 's, (0.10, 0.16, 0.05, 0.03, 0.34), reveals that the s.d. for component 5 is reduced. To understand this odd phenomenon, consider what happens when only component 5 varies, say by an amount Δu_5^* . In this case, $x_5 = 91.5(605.7 + \Delta u_5^*) / (915 + \Delta u_5^*)$ and the derivative with respect to Δu_5^* is approximately 0.034. If only component 2 varies, then $x_5 = 91.5(605.7) / (915 + \Delta u_2^*)$ with associate derivative approximately -0.066 . Using these as a basis for first-order Taylor series expansions for the standard deviation of x_5 , we see that this odd phenomenon will occur whenever $\sigma_2/\sigma_5 > 0.034/0.066 = 0.52$ —it is 0.60 for the example.

Returning to the example, we find that the total s.d. of Y is estimated to be 0.0333 with estimated percent variance contributions due to components A – E of (8, 16, 4, 40, 32). The percent contribution due to the non-additive portion was only 0.04%. Any reductions in variance that are needed must likely come from improving the precision of components D and E . For example, if $\sigma_4 = 0$, we find that the estimated variance of Y become 0.258 and $(.0258/.0333)^2 = 60\%$. The induced-variance effects noted in the previous paragraph have already been taken into account in these calculations, so the usual additive sums of variances continues to hold here, aside from the minor feature captured in the non-additive portion of the total variance.

11 Conclusions

We believe that our extension of mean-modeled, first-order, analysis of spread. to an analysis that is exact for quadratic process models and nearly so for quadratic mixture models should be useful in many applications. The ability to perform LANOVA—obtain functions that perform variance decomposition in a closed form—should also be valuable for such models. These techniques can also be useful for decomposing test uncertainty for deterministic models into factors as well as variance/bias decompositions. Examples have been provided to suggest the value of the method.

References

- ASME PRC 19.1-1998 (1998), *Test Uncertainty*, American Society of Mechanical Engineers.
- ASTM E 1169-02 (2002), “Standard Guide for Conducting Ruggedness Tests,” ASTM International.
- Box, G. E. P. (1954), “The Exploration and Exploitation of Response Surfaces: Some General Considerations and Examples,” *Biometrics*, 10, 16–60.
- Box, G. E. P. (1988), “Signal-to-Noise Ratios, Performance Criteria, and Transformations” (with discussion), *Technometrics*, 30, 1–40.
- Box, G. E. P., and Cox, D. R. (1964), “An Analysis of Transformations,” *Journal of the Royal Statistical Society B*, 26, 211–252.
- Box, G. E. P., and Draper, N. R. (1975), *Empirical Model Building and Response Surfaces*, John Wiley & Sons, New York.
- Box, G. E. P., and Meyer, R. D. (1986), “Dispersion Effects from Fractional Designs,” *Technometrics*, 28, 19–27.
- Derringer, G., and Suich, R., (1980), “Simultaneous Optimization of Several Response Variables,” *Journal of Quality Technology*, 12, 214–219. .
- Hamada, M. S., Martz, H. F., Steiner, S., (2005), “Accounting for Mixing Errors in Analyzing Mixture Experiments,” *Journal of Quality Technology*, 37, 139–148.
- ISO (1993), *Guide to the Expression of Uncertainty in Measurement*, International Standards Organization.
- León, R. V., Shoemaker, A. C., and Kacker, R. N. (1987), “Performance Measures Independent of Adjustment” (with discussion), *Technometrics*, 29, 253–285.
- Lucas, J. M. (1994), “How to Achieve a Robust Process Using Response Surface Methodology,” *Journal of Quality Technology*, 26, 248–260.
- Mandel, J. (1984), “Fitting Straight Lines When Both Variables are Subject to Error,” *Journal of Quality Technology*, 16, 1–14.
- Myers, R. H., Khuri, A. I., and Vining, G. (1992), “Response Surface Alternatives to the Taguchi’s Robust Parameter Design Approach,” *American Statistician*, 46, 131–139.

- Nair and Pregibon (1988), "Analyzing Dispersion Effects From Replicated Factorial Designs," *Technometrics*, 30, 247–257 (Corr: *Technometrics*, 31, 133).
- Paule, R. C. (1988), "Ruggedness tests and interlaboratory studies," *ASTM Standardization News*, 16, 56–58.
- Plante, R. (2002), "Multivariate tolerance design for a quadratic design parameter model," *IIE Transactions*, 34, 565–571.
- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989), "Design and analysis of computer experiments," *Statistical Science*, 4, 409–423 (C/R: 423–435).
- Searle, S. R. (1971), *Linear Models*, John Wiley, NY.
- Shoemaker, A. C., Tsui, K. L., and Wu, C. F. J. (1991), "Economical Experimental Methods for Robust Parameter Design," *Technometrics*, 33, 415–427.
- Steiner, S. H., and Hamada, M. (1997), "Making Mixtures Robust with Respect to Noise and Mixing Measurement Errors," *Journal of Quality Technology*, 29, 441–450.
- Taguchi (1987), *Systems of Experimental Design*, Vol 1 & 2, UNIPUB, White Plains, NY.
- Taylor, W. (1991), "Variation transmission analysis," *ASQC Quality Congress Transactions*, 653–659
- Vining, G. G., and Myers, R. H. (1990), "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach," *Journal of Quality Technology*, 22, 38–45.
- Welch, W, Yu, T. K, Kang, S. M., and Sacks, J. (1990), "Computer Experiments for Quality Control by Parameter Design," *Journal of Quality Technology*, 22, 15–33.
- Wernimont, G. (1977), "Ruggedness evaluation of test procedures," *ASTM Standardization News*, 5, 13-16
- Youden, W. J. (1961), "Experimental Design and ASTM Committees," *Materials Research and Standards*, 1, 862–867.