TECHNOMETRICS VOL. 13, NO. 1 FEBRUARY, 1971

A New Analysis of Variance Model for Non-additive Data

JOHN MANDEL

Institute for Materials Research, National Bureau of Standards, Washington, D. C.

INTRODUCTION

The simplest structure that a two-way table of responses z_{ij} can display is that of complete additivity. For this structure the model is

$$z_{ij} = A_i + B_j + \text{random error} \begin{cases} i = 1 \text{ to } m \\ j = 1 \text{ to } n \end{cases} \tag{1}$$

or

$$E(z_{ij}) = A_i + B_j \,. \tag{2}$$

Then z_{ij} is, apart from random error, partitionable into two additive parts. The simplicity and practical usefulness of this model, whenever it is applicable, results from the reduction of a function of two variables, z_{ij}, into two functions of a single variable each, A_i and B_j.

Using the common analysis of variance notation, the additive model is usually represented by

$$z_{ij} = \mu + \rho_i + \gamma_j + \epsilon_{ij} \,, \tag{3}$$

which differs from (2) only in that a general mean, μ, has been extracted from the data prior to the partitioning into row and column effects. The ϵ_{ij} term represents random error.

Few sets of data obey a strictly additive model. To obtain an algebraic representation for the more general, non-additive situation, it is customary to write

$$z_{ij} = \mu + \rho_i + \gamma_j + \eta_{ij} \,, \tag{4}$$

where η_{ij} is referred to as the interaction between rows and columns, and where η_{ij}, unlike ϵ_{ij}, is no longer considered as just random error. However, a *part* of η_{ij} may be just random error. Because of the presence of η_{ij}, which, apart from its random error component, is a function of two variables, all advantage of the additive model will be lost, unless one can again partition the non-random portion of η_{ij} into functions of only one variable each. Obviously, an *additive* partitioning of η_{ij} is impossible since all additive parts have already

Received Jan. 1969; revised May 1969.

been extracted from the model. We propose a partitioning of η_{ij} into the sum of *multiplicative* functions of i and j, according to the model

$$z_{ij} = (\mu + \rho_i + \gamma_j) + (\theta_1 u_{1i} v_{1j} + \theta_2 u_{2i} v_{2j} + \cdots). \tag{5}$$

Thus, the interaction term is represented by

$$\eta_{ij} = \sum_{k=1}^{K} \theta_k u_{ki} v_{kj}. \tag{6}$$

As mentioned above, part of the interaction may well be random error. Consequently, in practice, the partitioning of the interaction into multiplicative terms is carried out only partially; i.e., only a few multiplicative terms of the $\theta u_i v_j$ type (generally one or two such terms) are retained; the remaining terms are pooled together and considered as experimental error. The theory of this partitioning as well as practical details of the procedure will be taken up after discussion of the illustrative example.

A Numerical Illustration

The data in Table EI are measurements of the density of aqueous solutions of ethyl alcohol of different concentrations, at temperatures ranging from 10 to 40°C [6]. Our aim is twofold: 1) to learn as much as possible about the structure of these data; and 2) to express density as an explicit function of concentration and temperature. The function will be empirical, but it should be valid over the entire range of the independent variables covered by the data.

Table EI

Density of Aqueous Solutions of Ethyl Alcohol

Concentration	Temperature						
	10.	15.	20.	25.	30.	35.	40.
30.086	.959652	.956724	.953692	.950528	.947259	.943874	.940390
39.988	.942415	.938851	.935219	.931502	.927727	.923876	.919946
49.961	.921704	.917847	.913922	.909938	.905880	.901784	.897588
59.976	.899323	.895289	.891202	.887049	.882842	.878570	.874233
70.012	.875989	.871848	.867640	.863378	.859060	.854683	.850240
80.036	.851882	.847642	.843363	.839030	.834646	.830202	.825694

The first step in our solution is to make an analysis of variance as shown in Table EII. It will be noted that in addition to the usual three terms (the two main effects and the interaction between rows and columns) there appear several more items, corresponding to the partitioning of the interaction term. The additional terms follow the usual pattern of the analysis of variance: each one involves a number of degrees of freedom (fractional in this case), a sum of squares, and a mean square. The sums of squares θ_1^2 , θ_2^2 , and θ_3^2 are computed as explained in the Appendix, and the corresponding degrees of freedom are taken from Table II, using $r = m - 1 = 5$, and $s = n - 1 = 6$. The sum of squares and the degrees of freedom for "residual" (last line) are obtained by difference. All mean squares are obtained in the usual manner.

TABLE EII

Density of Alcohol Analysis of Variance

Term in Model	Degrees of Freedom	Sum of Squares	Mean Square
μ	1	33.775858	33.775858
ρ_i	5	.061588	.012317
γ_j	6	.002646	.000441
η_{ij}	30	26.111383 x 10^{-6}	870,000 x 10^{-12}
$\theta_1 u_{1i} v_{1j}$	15.5	26.107899 "	1,690,000 "
$\theta_2 u_{2i} v_{2j}$	8.2	0.003268 "	450 "
$\theta_3 u_{3i} v_{3j}$	4.2	0.000148 "	35 "
Residual	2.1	0.000069 "	33 "

Residual variance (2 last lines combined) = 34×10^{-12} $\sigma_\epsilon = 6 \times 10^{-6}$

In the present case, the interaction has been partitioned into three multiplicate terms and a remainder (residual). The mean square for the third term is practically identical with that for the residual, and is therefore pooled with that term to give an estimate for the error variance.* The first two multiplicative terms, on the other hand, and especially the first of these have mean squares that are considerably larger than the pooled error variance. They represent real effects.

It may be concluded that the following model represents the data to within their experimental error.

$$D_{ij} = \mu + \rho_i + \gamma_j + \theta_1 u_i v_j + \theta_2 u_i' v_j' + \epsilon_{ij}$$

or**:

$$D_{ij} = \mu + Rr_i + Gc_j + \theta_1 u_i v_j + \theta_2 u_i' v_j' + \epsilon_{ij} \tag{E1}$$

with $\sigma_\epsilon = 6 \times 10^{-6}$.

The next step in the analysis involves an examination of the parameters occurring in this equation, particularly of the vectors r_i and c_j and the "eigenvectors" u_i , v_j , u_i' , v_j' , both in their relationship to each other and to the independent variables. We will refer to all parameters occurring in Eqs. (5) and (E1) as "structural parameters"; their numerical values, for our example, are shown in Table EIII.

* The standard deviation of error thus calculated is 6×10^{-6}. The data are recorded to six decimals, but the last decimal is of doubtful validity.

** The "eigenvectors" u_i , u_i' , v_j , v_j' are "normalized" in the sense that $\sum u_i^2 = \sum u_i'^2 = \sum v_j^2 = \sum v_j'^2 = 1$. For purposes that will soon become apparent, it is advantageous to also normalize the vectors ρ_i and γ_j of Eq. (5) (see Eqs. (11) and (12)).

TABLE EIII

Density of Alcohol—Values of Structural Parameters

a. Constants

$\hat{\mu}$ = .896,764,8; R = .093,799,1; G = .021,000,6

$\hat{\theta}_1$ = .005,109,6 $\hat{\theta}_2$ = .000,057,2

b. Row Parameters

Concentration	r_i	u_i	u_i'
30.086	.570772	.784749	-.459360
39.988	.368846	.231494	.724416
49.961	.139065	-.052135	.329347
59.976	-.104853	-.220479	-.049510
70.012	-.357170	-.334699	-.321264
80.036	-.616659	-.408933	-.223783

c. Column Parameters

Temperature	c_j	v_j	v_j'
10	.558524	-.590282	.511802
15	.377862	-.377433	.008420
20	.194035	-.173152	-.337774
25	.006635	.018319	-.468378
30	-.183922	.201870	-.355629
35	-.377766	.376021	.129797
40	-.575363	.544656	.511083

It should be noted that whereas u_i is a monotonic function of r_i , and v_j is monotonic in c_j , the parameters u_i' and v_j' pass through a maximum and minimum, respectively. We will see that according to theory u_i and u_i' are orthogonal, i.e., *linearly* independent, but they are obviously not *functionally* independent. The behavior of u_i' suggests that it may be expressible as a quadratic function of u_i . Consider the function $u_i^* = au_i^2 + bu_i + c$, where $\sum_i u_i = 0$, $\sum_i u_i^2 = 1$, $i = 1$ to m, and impose the conditions:

$$\sum u_i^* = 0$$

$$\sum u_i^* u_i = 0$$

$$\sum u_i^{*2} = 1.$$

Then it follows that

$$u_i^* = C\left[u_i^2 - (\sum u_i^3)u_i - \frac{1}{m}\right]$$

where the constant C is such that $\sum u_i^{*2} = 1$. This yields two solutions for C, say C and $(-C)$. Thus, the normalized form of any quadratic function of u_i , if it is to be orthogonal to u_i , must be either u_i^* or $(-u_i^*)$. Table EIV shows both u_i' and u_i^* , and it is seen that, to within approximately ± 0.1, we have $u_i' = -u_i^*$. If we substitute $(-u_i^*)$ for u_i' , the error induced in D_{ij} (Eq. (E1)) will be less than (θ_2) (0.1) or 6×10^{-6}. Thus, this substitution is empirically ac-

ceptable. A similar procedure leads us to substitute v_j^* for v_j' (see Table EIV), where v_j^* is derived from v_j as u_i^* is derived from u_i . Table EIV also lists the coefficients necessary for the calculation of u_i^* and v_j^* .

TABLE EIV

Relations Between First and Second Eigenvectors

u_i^*	u_i'	v_j^*	v_j'
.4487	-.4594	.5516	.5118
-.5928	.7244	-.0491	.0084
-.4257	.3293	-.3660	-.3378
-.1019	-.0495	-.4320	-.4684
.2131	-.3213	-.2856	-.3556
.4591	-.2238	.0433	.1298
		.5379	.5111

$u_i^* = 2.95216\ u_i^3 - 1.11869\ u_i - 0.492027$

$v_j^* = 3.04771\ v_j^3 + 0.126972\ v_j - 0.435387$

The only remaining task is to express the quantities r_i and u_i as functions of the concentration, and c_j and v_j as functions of the temperature. Before completing the empirical fit in this way, however, it is interesting to observe the remarkable similarity (except for an inversion of all signs) between the vectors c_j and v_j . If we identified v_j with $(-c_j)$, the empirical fit would suffer somewhat, but it is interesting, nevertheless, to examine the consequences of such an approximation in terms of the mathematical model.

Let us then assume, tentatively, that $v_j = -c_j$ and let us, at the same time, neglect the term $\theta_2 u_i' v_j'$ (which is only of the order of 6×10^{-5}). Under these assumptions, the model (E1) becomes

$$D_{ij} = \mu + Rr_i + (G - \theta u_i)c_j + \epsilon_{ij} , \tag{E2}$$

which may be rewritten in the form

$$D_{ij} = \mu + \rho_i + \beta_i\gamma_j + \epsilon_{ij} . \tag{E3}$$

This is a model of the general type

$$D_{ij} = f(i) + g(i)h(j) + \epsilon_{ij} , \tag{E4}$$

which has been discussed by the author in a previous paper [2].

If we fitted this model to our data, we would obtain a residual standard deviation (estimated with 25 degrees of freedom) of 5×10^{-5}. Furthermore, the vector β_i , of Eq. (E3), when normalized, would become

$$[-.784443,\ -.231949,\ .051948,\ .220520,\ .334918,\ .409085]$$

which, apart from a reversal of all signs, is practically identical with the vector u_i of Table EIII.

TABLE EV

Density of Alcohol

Formulas for Empirical Fit

$$z_{ij} = 0.896765 + 0.093799\, r_i + 0.021001\, c_j + 0.005{,}109\, u_i v_j + 0.000057\, u_i' v_j'$$

$$\begin{cases} r_i = 0.457609 - 0.018006\,(x_i - 35.589532)\,|x_i - 35.589532|^{0.077496} \\ u_i = -0.784676 + 107.858450\,(x_i - 4.882253)^{-1.310860} \\ u_i' = -u_i^{*} \text{ (see Table EIV)} \end{cases}$$

$$\begin{cases} c_j = 5.289{,}976 - 4.380455\,\left(1 + \dfrac{y_j}{252.919650}\right)^{1.988011} \\ v_j = -15.778065 + 6.620529\,(y_j + 57.866579)^{0.196868} \\ v_j' = v_j^{*} \text{ (see Table EIV)} \end{cases}$$

x_i = concentration $\qquad$ y_j = temperature

Use of the slightly more complicated model (E1) will result in a substantial improvement in the accuracy of fit. To complete the analysis we must fit the vectors r_i and u_i to "concentration", and c_j and v_j to "temperature". To this end, use was made of the general formula*

$$y = y_0 + A(x - x_0)\,|x - x_0|^{B-1}. \tag{E5}$$

The four constants y_0, A, x_0, and B were calculated by an iterative least squares procedure. The results are given in Table EV which lists all numerical equations required for the fit. The "percent residuals" of this fit, i.e., the quantities

$$100\,\frac{\text{fitted value-observed value}}{\text{observed value}}$$

are shown in Table EVI. It is seen that the accuracy of fit is of the order of 0.004%. No percent residual exceeds 0.007% and no residual exceeds 0.00006 density units. Had we used model (E3) we would have obtained residuals about ten times larger than those shown in Table EVI.

An important use of an empirical fit such as the one just discussed, lies in its role as an interpolation device. The formula provides values of the response for any combination of values of the independent variables, provided that the latter are inside the range covered by the experimental data. Since all functions used in the empirical fit are monotonic, the fit may be expected to have about the same precision for interpolated values as it has for directly fitted values. While interpolation is greatly facilitated by an empirical formula, any attempt to use it for extrapolation outside the range of values covered by the experimental data should be avoided or undertaken only with the greatest caution, even when the extrapolation is to a region close to that covered by the data.

* In many cases, one can use the simpler formulas $y = y_0 + A(x - x_0)^B$ or $y = y_0 + A(x_0 - x)^B$. Eq. (E5) contains both of these forms and is applicable even when x_0 is inside the range of the x values.

TABLE EVI

Density of Alcohol—Percent Residuals[1] of Empirical Fit

Concentration	Temperature 10.	15.	20.	25.	30.	35.	40.
30.086	.00089672	-.00011837	.00096014	.00054790	.00092261	.00046336	.00057124
39.988	.00118904	.00064836	.00078472	-.00023276	.00047624	.00097419	.00091275
49.961	-.00463022	-.00391099	-.00401094	-.00380333	-.00481802	-.00297186	-.00505511
59.976	-.00120045	-.00078892	-.00004431	-.00024190	.00012828	-.00004240	-.00064941
70.012	.00649297	.00738266	.00683711	.00653345	.00641104	.00626342	.00538656
80.036	-.00330687	-.00518157	-.00527765	-.00516904	-.00432763	-.00369028	-.00362110

[1] Percent residual = $100 \frac{\text{fitted value} - \text{observed value}}{\text{Observed value}}$

THEORETICAL RESULTS[1]

We now return to Eqs. (5) and (6), and concern ourselves with the estimation of the quantities θ_k , u_{ki} , and v_{kj} . The estimation is accomplished by application of the method of least squares. The details have been developed in a previous paper [4] in which the following facts are also derived.

1. There are at most $(m - 1)$ or $(n - 1)$—whichever is the smaller—number of terms in the partitioning of η_{ij} according to Eq. (6). Thus, $K \leq \min (m - 1, n - 1)$. It follows that a *finite* set of multiplicative terms, at most $(m - 1)$ or $(n - 1)$ in number, suffices to express the total interaction between rows and columns.

2. The parameters θ_k , u_{ki} , v_{kj} , for *all* values of k, are estimated entirely from the matrix of residuals d_{ij} , where

$$d_{ij} = z_{ij} - \hat{\mu} - \hat{\rho}_i - \hat{\gamma}_j = z_{ij} - z_{..} - (z_{i.} - z_{..}) - (z_{.j} - z_{..}), \qquad (7)$$

(where a dot indicates averaging over the index replaced by a dot). This is the usual matrix of residuals in the analysis of variance, resulting from the elimination of row effects, column effects, and of the grand mean.

In fact, the procedure is sequential, so that any term $\theta_k u_{ki} v_{kj}$, and all subsequent terms, are completely determined by the table of residuals obtained by stopping after the term $\theta_{k-1} u_{k-1,i} v_{k-1,j}$.

3. Denoting estimates by a caret ($\hat{}$) placed over the symbols of the parameters $(\hat{\theta}, \hat{u}, \hat{v})$, we have

$$\sum_i \hat{u}_{ki} = \sum_j \hat{v}_{kj} = 0, \qquad \text{for all } k \qquad (8)$$

$$\sum_i \hat{u}_{ki}^2 = \sum_j \hat{v}_{kj}^2 = 1, \qquad \text{for all } k. \qquad (9)$$

By definition, each $\hat{\theta}_k$ is positive (unless it is zero).

Equations (8) and (9) are expressed by stating that the vectors $\hat{u}_{ki}$ and $\hat{v}_{kj}$ are "normalized."

4. For any two values of k, say k and k' $(k \neq k')$ we have

[1] The understanding of this section will be facilitated if the reader will refer back to the illustrative example wherever applicable.

$$\sum_i u_{ki}u_{k'i} = 0$$
$$\sum_j v_{kj}v_{k'j} = 0 \tag{10}$$

Equations (10) are expressed by stating that the vectors $\hat{u}_{ki}$ and $\hat{u}_{k'i}$ are *orthogonal*, with a similar statement for v_{kj} and $v_{k'j}$.

For reasons that were discussed in connection with the illustrative example, it is advantageous to also "normalize" the vectors $\hat{\rho}_i$ and $\hat{\gamma}_j$. The first condition $\sum_i \hat{\rho}_i = \sum_j \hat{\gamma}_j = 0$ is always fulfilled. In order to satisfy the second condition of normalization, we write

$$\hat{\rho}_i = Rr_i \qquad \text{and} \qquad \hat{\gamma}_j = Gc_j \,, \tag{11}$$

where

$$R = \sqrt{\sum_i \hat{\rho}_i^2} \quad \text{and} \quad G = \sqrt{\sum_j \hat{\gamma}_j^2} \,. \tag{12}$$

It then follows that

$$\sum_i r_i = \sum_j c_j = 0 \tag{13}$$

$$\sum_i r_i^2 = \sum_j c_j^2 = 1. \tag{14}$$

Corresponding to Eq. (5), we now can write a relation in terms of parameter estimates, as follows:

$$z_{ij} = (\hat{\mu} + Rr_i + Gc_j) + \sum_{k=1}^{K} \hat{\theta}_k \hat{u}_{ki}\hat{v}_{kj} \,. \tag{15}$$

The analog of Eq. (6) becomes:

$$d_{ij} = \sum_{k=1}^{K} \hat{\theta}_k \hat{u}_{ki}\hat{v}_{kj} \tag{16}$$

since d_{ij} , as given by Eq. (7), is the least squares estimate of η_{ij} .

It has also been shown in the previous paper [4] that to the partitioning of the interaction according to Eq. (16), there corresponds a parallel partitioning of the *sum of squares of interaction* of the usual analysis of variance table. This partitioning is based on the identity

$$\sum_i \sum_j d_{ij}^2 = \sum_{k=1}^{K} \hat{\theta}_k^2 \,, \tag{17}$$

where the components of the sum of squares, $\hat{\theta}_k^2$, are simply the squares of the estimates of the coefficients θ_k of Eq. (6).

As noted above, the main advantage of the additive model, i.e., the possibility of expressing a function of two variables in terms of functions of a single variable, is retained in the more general model (5), since all u terms are functions of i only and all v terms functions of j only.

If the partitioning is complete, using all K terms in Eqs. (6) and (16), the residuals remaining at the end of this process will all be zero, leaving no estimate for the experimental error ϵ_{ij} . In practice, one generally stops the par-

titioning process long before this has occurred. For example, if the partitioning of the interaction is limited to two terms, then the remaining $(K - 2)$ terms provide an estimate of the residual error ϵ_{ij} . The main problem is to derive a criterion by which a decision can be made as to where the partitioning process should be ended. After this decision has been made, using criteria to be developed in the following two sections, the next phase of the analysis consists in examining:

a) the relationships between the vectors r_i , u_i , u'_i , etc., with each other* and with the independent variable (or combination of variables) representing the rows (the "row labels"), and

b) the relationships between the vectors c_j , v_j , v'_j , etc., with each other* and with the "column labels".

This second phase generally results in a considerably simpler model, requiring fewer constant and vectors, for the functional representation of the data.

Both of these phases of the analysis were discussed in detail in connection with the numerical illustrative example.

The Analysis of Variance Table

The additive model represented by Eq. (3) is usually represented by an analysis of variance table containing sums of squares, degrees of freedom, and mean squares for the grand mean μ, the row effects ρ_i , the column effects γ_j , and for the "residuals" ϵ_{ij} .

If we are to represent the non-additive model of Eqs. (5) and (15) in an analogous way, we must find, for each term $\hat{\theta}\hat{u}_i\hat{v}_j$: a corresponding sum of squares, a number of degrees of freedom, and a mean square. A schematic of such a partitioning process is shown in Table I.

The sums of squares corresponding to the multiplicative terms are, as mentioned above, simply the quantities $\hat{\theta}_1^2$, $\hat{\theta}_2^2$, $\cdots$, $\hat{\theta}_k^2$, $\cdots$ and these can be computed (see Appendix A) by appropriate methods. On the other hand, there is no simple way by which the "degrees of freedom" ν_1 , ν_2 , etc., can be written down in analogy to the degrees of freedom of ordinary analysis of variance tables.** The reason for this is that whereas the quantities $\hat{\theta}_k^2$ are legitimate additive components of the sum of squares $\sum_i \sum_j d_{ij}^2$, they are *not* distributed in accordance with the chi-square distribution. We must first *define* degrees of freedom in a more general way, before attempting to calculate them.

Degrees of Freedom

The clue to an appropriate definition of degrees of freedom in the non-additive case is given by the entries in the last column of Table I. The number of degrees

* The vectors u_i, u'_i, etc. are orthogonal and therefore *linearly* independent. They may, however, be, and often are, functionally dependent. Furthermore, r_i is generally not even linearly independent of the u_i, u'_i, etc. A similar situation holds for the column dependent vectors c_j, v_j, v'_j, etc.

** The various attempts to do this, as discussed in the literature, are all erroneous (see Ref. [4]).

TABLE I

Analysis of Variance for Non-additive Model

Term in Model (Eq.5)	Degrees of Freedom	Sum of Squares	Mean Square
μ	1	Usual formulas	
ρ_i	m-1		
γ_j	n-1		
η_{ij}	(m-1)(n-1)		
$\theta_1 u_{1i} v_{1j}$	ν_1	$\hat{\theta}_1^2$	$\hat{\theta}_1^2/\nu_1$
$\theta_2 u_{2i} v_{2j}$	ν_2	$\hat{\theta}_2^2$	$\hat{\theta}_2^2/\nu_2$
$\theta_k u_{ki} v_{kj}$	ν_k	$\hat{\theta}_k^2$	$\hat{\theta}_k^2/\nu_k$

of freedom ν_k is the *divisor* corresponding to $\hat{\theta}_k^2$. It must be such that the resulting quotient (which we still call mean square, in analogy to the additive case) is, *in the absence of real systematic interaction effects*, merely an estimate of random experimental error. Thus, if σ^2 is the variance of the random error ϵ_{ij}, and if none of the terms $\theta_k u_{ki} v_{kj}$ occur in the real model, then we must have

$$E(\hat{\theta}_k^2/\nu_k) = \sigma^2 \qquad \text{for each } k. \tag{18}$$

Equation (18) can be written

$$\nu_k = E(\hat{\theta}_k^2)/\sigma^2. \tag{19}$$

Equation (19) *provides our definition for the degrees of freedom.* It also provides us with a way of calculating the ν_k, as explained in the following section.

CALCULATION OF DEGREES OF FREEDOM

Suppose we start with an $m \times n$ matrix of *independent, normally distributed* variates z_{ij}, of common variance σ^2, and assume that

$$E(z_{ij}) = \mu + \rho_i + \gamma_j . \tag{20}$$

These z_{ij} obey a strictly additive model. If we analyze them, nevertheless, in accordance with the analysis of variance shown in Table I, then the partitioning of the interaction should faithfully reflect the absence of any real terms of the type $\theta u_i v_j$ in the model. Thus, each of the mean squares $\hat{\theta}_k^2/\nu_k$ must then simply be an estimate of σ^2, in accordance with Eq. (18).

For convenience we can take our z_{ij} from a table of random normal deviates. Then $\sigma^2 = 1$ and $\mu = \rho_i = \gamma_j = 0$ for all i and j. In that case, it follows from Eq. (19) that

$$\nu_k = E(\hat{\theta}_k^2). \tag{21}$$

Thus, in order to obtain ν_k it is only necessary to repeat the analysis a large number of times, taking a different set of random normal deviates each time and averaging, for each value of k separately, the $\hat{\theta}_k^2$ values obtained in all sets.

A Monte Carlo experiment of the type just described was carried out on a high-speed computer, using 625 independent sets of z_{ij} for each combination of m and n, with:

$$m = 4(1)8(2)12, 16, 20, 32, 50, 100$$

$$n = 4(1)8, 10, 12, 16, 20.$$

In addition to the average of the $\hat{\theta}_k^2$, for each value of k, we also calculated their standard deviations. The averages and standard deviations are shown in Table II and III, respectively. Rather than labelling the rows and columns

TABLE II

Values of ν_k for k = 1, 2, 3

k	r\s	2	3	4	5	6	7	9	11	15	19
1	2	3.55	5.03	6.34	7.53	8.96	10.15	12.77	15.09	19.58	24.10
	3	5.03	6.71	8.33	9.85	11.32	12.74	15.56	18.08	23.23	28.34
	4	6.34	8.33	10.21	11.81	13.40	15.03	18.16	21.00	26.51	31.83
	5	7.53	9.85	11.81	13.51	15.47	16.97	20.23	23.24	29.14	34.90
	6	8.96	11.32	13.40	15.47	17.17	19.11	22.40	25.63	31.56	37.58
	7	10.15	12.74	15.03	16.97	19.11	20.96	24.35	27.77	34.24	40.30
	9	12.77	15.56	18.16	20.23	22.34	24.35	27.99	32.04	38.79	45.38
	11	15.09	18.08	21.00	23.24	25.63	27.77	32.04	36.25	42.73	50.33
	15	19.58	23.23	26.51	29.14	31.56	34.24	38.79	42.73	50.84	58.76
	19	24.10	28.34	31.83	34.90	37.58	40.30	45.38	50.33	58.76	65.94
	31	37.88	42.87	47.26	50.70	54.71	57.34	63.52	68.65	78.89	88.15
	49	57.85	63.78	69.31	73.46	77.51	81.39	88.82	95.34	106.82	117.83
	99	111.22	120.09	127.27	133.44	138.99	144.25	152.80	161.46	176.77	192.05
2	2	0.45	0.97	1.66	2.47	3.04	3.85	5.23	6.91	10.42	13.90
	3	0.97	2.01	3.01	4.00	4.98	6.01	7.98	10.02	14.04	17.96
	4	1.66	3.01	4.21	5.45	6.66	7.88	10.12	12.50	16.90	21.25
	5	2.47	4.00	5.45	7.02	8.22	9.70	12.08	14.77	19.62	24.28
	6	3.04	4.98	6.66	8.22	9.76	11.28	14.06	16.79	22.05	27.10
	7	3.85	6.01	7.88	9.70	11.28	12.82	15.78	18.76	24.14	29.59
	9	5.23	7.98	10.12	12.08	14.06	15.78	19.30	22.18	28.59	34.00
	11	6.91	10.02	12.50	14.77	16.79	18.76	22.18	25.60	32.42	38.45
	15	10.42	14.04	16.90	19.62	22.05	24.14	28.59	32.42	39.55	46.40
	19	13.90	17.96	21.25	24.28	27.10	29.59	34.00	38.45	46.40	53.54
	31	24.12	29.99	34.26	38.32	41.60	44.55	50.38	55.57	65.47	74.01
	49	40.15	48.10	53.79	58.20	62.55	66.40	73.25	80.24	90.97	101.74
	99	86.78	97.81	106.11	112.48	118.85	123.83	133.46	142.55	157.22	171.24
3	3		0.28	.66	1.15	1.69	2.25	3.46	4.89	7.72	10.70
	4		0.66	1.37	2.22	3.06	3.76	5.43	7.11	10.70	14.33
	5		1.15	2.22	3.19	4.16	5.24	7.24	9.22	13.24	17.17
	6		1.69	3.06	4.16	5.44	6.51	8.85	11.06	15.47	19.73
	7		2.25	3.76	5.24	6.51	7.81	10.44	12.82	17.55	22.24
	9		3.46	5.43	7.24	8.85	10.44	13.43	11.07	21.32	26.37
	11		4.89	7.11	9.22	11.06	12.82	16.07	19.07	25.02	30.58
	15		7.72	10.70	13.24	15.47	17.55	21.32	25.02	31.71	37.95
	19		10.70	14.33	17.17	19.73	22.24	26.37	30.58	37.95	44.89
	31		20.14	25.27	29.30	32.66	35.60	41.34	46.33	55.94	64.20
	49		35.12	41.77	47.46	51.66	55.60	62.55	68.84	79.94	90.41
	99		79.09	89.71	97.33	103.25	109.53	119.41	127.69	143.41	156.36

m and n, of these tables (as was done in Ref. [4]), we have labelled them r and s for reasons which will be explained in the next section.

The values in Table II differ slightly from the corresponding values listed in Table A1 of Reference [4]. In the present paper, the values were corrected by a least squares adjustment procedure, and are consistent with those listed in Table IV, and with Table A3 of Ref. [4].

TABLE III

Standard Deviations of $\hat{\theta}_k^2$ for $k = 1, 2, 3$

k	r\s	2	3	4	5	6	7	9	11	15	19
1	2	2.68	3.08	3.16	3.56	3.92	4.01	4.64	5.00	5.65	6.05
	3	3.08	3.24	3.83	4.05	4.36	4.52	4.72	5.26	5.71	6.25
	4	3.16	3.83	4.18	4.12	4.28	4.71	5.12	5.45	5.84	6.40
	5	3.56	4.05	4.12	4.30	4.88	4.71	5.64	5.51	6.09	6.92
	6	3.92	4.36	4.28	4.88	4.80	5.16	5.28	5.64	6.03	6.51
	7	4.01	4.52	4.71	4.71	5.16	5.26	5.30	6.02	6.49	6.88
	9	4.64	4.72	5.12	5.64	5.28	5.30	5.50	6.41	6.25	6.89
	11	5.00	5.26	5.45	5.51	5.64	6.02	6.41	6.65	6.57	7.53
	15	5.65	5.71	5.84	6.09	6.03	6.49	6.25	6.57	7.10	7.60
	19	6.05	6.25	6.40	6.92	6.51	6.88	6.89	7.53	7.60	7.87
	31	7.22	7.38	7.59	7.96	7.89	7.98	8.02	8.26	8.37	8.77
	49	9.07	9.55	9.44	8.88	9.04	9.36	9.77	9.86	10.16	
	99	12.98	12.96	12.22	12.50	11.84	11.96	11.67	11.59	11.95	12.19
2	2	0.66	0.98	1.38	1.69	1.87	2.00	2.50	2.87	3.71	4.39
	3	0.98	1.39	1.77	2.08	2.25	2.62	2.77	3.17	3.56	4.11
	4	1.38	1.77	2.04	2.29	2.48	2.70	3.11	3.43	3.95	4.32
	5	1.69	2.08	2.29	2.61	2.83	2.94	3.36	3.68	4.00	4.46
	6	1.87	2.25	2.48	2.83	3.04	3.04	3.36	3.77	4.20	4.72
	7	2.00	2.62	2.70	2.94	3.04	3.17	3.43	3.82	4.26	4.75
	9	2.50	2.77	3.11	3.36	3.36	3.43	3.99	4.09	4.66	5.00
	11	2.87	3.17	3.43	3.68	3.77	3.82	4.09	4.10	4.85	4.93
	15	3.71	3.56	3.95	4.00	4.20	4.26	4.66	4.85	5.30	5.55
	19	4.39	4.11	4.32	4.46	4.72	4.75	5.00	4.93	5.55	5.83
	31	5.49	5.48	5.24	5.60	5.44	5.52	5.82	5.88	6.27	6.20
	49	7.57	6.93	7.09	6.79	6.96	6.61	6.92	7.08	7.36	7.20
	99	10.34	10.15	9.59	9.08	9.56	9.04	9.23	8.90	9.27	9.41
3	3		.37	.69	.92	1.20	1.35	1.72	2.13	2.80	3.14
	4		.69	.97	1.23	1.56	1.66	1.86	2.30	2.90	3.36
	5		.92	1.23	1.53	1.75	2.00	2.21	2.46	3.08	3.44
	6		1.20	1.56	1.75	2.00	2.07	2.43	2.66	3.32	3.62
	7		1.35	1.66	2.00	2.07	2.31	2.62	2.80	3.19	3.56
	9		1.72	1.86	2.21	2.43	2.62	2.95	3.11	3.53	3.68
	11		2.13	2.30	2.46	2.66	2.80	3.11	3.34	3.60	3.87
	15		2.80	2.90	3.08	3.32	3.19	3.53	3.60	4.17	4.45
	19		3.14	3.36	3.44	3.62	3.56	3.68	3.87	4.45	4.76
	31		4.59	4.29	4.53	4.55	4.80	4.52	4.98	5.09	5.38
	49		6.34	5.64	5.68	5.67	5.84	5.78	5.77	5.87	6.01
	99		9.21	8.27	8.12	7.77	7.98	7.83	7.36	8.11	8.05

GENERALIZATION OF RESULTS

The partitioning of the sum of squares $\sum_i \sum_j d_{ij}^2$ can conceivably be carried out on sets of residuals obtained in a variety of ways. For example, instead of Eq. (5), we could have written

$$z_{ij} = (\mu + \rho_i) + (\theta_1 u_{1i} v_{ij} + \theta_2 u_{2i} v_{2j} + \cdots) + \epsilon_{ij} . \tag{22}$$

Equation (22) differs from Eq. (5) in that the term γ_j was not removed from the observations prior to partitioning the residuals. Thus, here the residuals that are partitioned are

$$z_{ij} - \hat{\mu} - \hat{\rho}_i .$$

Similarly, we could have partitioned the residuals

$$z_{ij} - \hat{\mu} - \hat{\gamma}_j ,$$

or even the original observations z_{ij} themselves.

More generally, we might have extracted from the original observations, prior to partitioning into sums of multiplicative terms, not only $\hat{\mu}$, $\hat{\rho}_i$, and $\hat{\gamma}_j$, but any number of "sweeps" of the "vacuum cleaner". [7]

It can be shown [5] that Tables II and III are applicable to all these cases, provided that the following rule is observed.

TABLE IV

ν_k expressed as % of total interaction degrees of freedom for k = 1, 2, 3

k	r\s	2	3	4	5	6	7	9	11	15	19
1	2	88.73	83.78	79.24	75.32	74.70	72.51	70.92	68.61	65.26	63.43
	3	83.78	74.54	69.42	65.66	62.91	60.68	57.64	54.80	51.63	49.72
	4	79.24	69.42	63.83	59.06	55.83	53.67	50.46	47.74	44.18	41.89
	5	75.32	65.66	59.06	54.05	51.57	48.48	44.96	42.26	38.85	36.74
	6	74.70	62.91	55.83	51.57	47.70	45.50	41.47	38.84	35.07	32.96
	7	72.51	60.68	53.67	48.48	45.50	42.79	38.66	36.07	32.61	30.30
	9	70.92	57.64	50.46	44.96	41.47	38.66	34.55	32.36	28.73	26.54
	11	68.61	54.80	47.74	42.26	38.84	36.07	32.36	29.96	25.90	24.08
	15	65.26	51.63	44.18	38.85	35.07	32.61	28.73	25.90	22.59	20.62
	19	63.43	49.72	41.89	36.74	32.96	30.30	26.54	24.08	20.62	18.27
	31	61.09	46.09	38.12	32.71	29.41	26.42	22.77	20.13	16.96	14.97
	49	59.03	43.39	35.36	29.98	26.36	23.73	20.14	17.69	14.53	12.66
	99	56.17	40.44	32.14	26.96	23.40	20.82	17.15	14.83	11.90	10.21
2	2	11.27	16.22	20.76	24.68	25.30	27.49	29.08	31.39	34.74	36.57
	3	16.22	22.39	25.06	26.69	27.67	28.61	29.56	30.38	31.21	31.51
	4	20.76	25.06	26.34	27.25	27.75	28.13	28.13	28.40	28.17	27.96
	5	24.68	26.69	27.25	28.09	27.41	27.72	26.85	26.86	26.17	25.55
	6	25.30	27.67	27.75	27.41	27.13	26.86	26.04	25.44	24.80	23.77
	7	27.49	28.61	28.13	27.72	26.86	26.17	25.05	24.36	22.99	22.24
	9	29.08	29.56	28.13	26.85	26.04	25.05	23.83	22.41	21.18	19.88
	11	31.39	30.38	28.41	26.86	25.44	24.36	22.41	21.16	19.65	18.40
	15	34.74	31.21	28.17	26.17	24.50	22.99	21.18	19.65	17.58	16.28
	19	36.57	31.51	27.96	25.55	23.77	22.24	19.88	18.40	16.28	14.83
	31	38.91	32.25	27.62	24.73	22.37	20.53	18.06	16.30	14.08	12.56
	49	40.97	32.72	27.44	23.76	21.28	19.36	16.61	14.89	12.38	10.93
	99	43.83	32.93	26.80	22.72	20.01	17.87	14.98	13.09	10.59	9.10
3	3		3.07	5.52	7.65	9.42	10.71	12.80	14.82	17.16	18.77
	4		5.52	8.56	11.08	12.75	13.41	15.09	16.15	17.83	18.86
	5		7.65	11.08	12.76	13.86	14.96	16.10	16.77	17.66	18.08
	6		9.42	12.75	13.86	15.12	15.50	16.38	16.76	17.19	17.30
	7		10.71	13.41	14.96	15.50	15.94	16.57	16.65	16.71	16.72
	9		12.80	15.09	16.10	16.38	16.57	16.58	16.23	15.79	15.42
	11		14.82	16.15	16.77	16.76	16.65	16.23	15.76	15.16	14.63
	15		17.16	17.83	17.66	17.19	16.71	15.79	15.16	14.09	13.32
	19		18.77	18.86	18.08	17.30	16.72	15.42	14.63	13.32	12.44
	31		21.66	20.38	18.90	17.56	16.41	14.82	13.59	12.03	10.90
	49		23.89	21.31	19.37	17.57	16.21	14.18	12.77	10.88	9.71
	99		26.63	22.65	19.66	17.38	15.81	13.40	11.73	9.66	8.31

Rule for Use of Tables II and III:

Tables II and III are applicable to any interaction matrix of r degrees of freedom by s degrees of freedom [5]. The parameter r may be thought of as a number of degrees of freedom associated with the rows of the table, and the parameter s as a number of degrees of freedom associated with the columns of the table. Tables II and III should be entered with values of r and s obtained as follows:

1. Partitioning on original observations z_{ij} (neither row nor column averages removed):

$$r = m, \qquad s = n.$$

2. Partitioning on $z_{ij} - \hat{\mu} - \hat{\rho}_i$ (row averages removed; column averages not removed):

$$r = m, \qquad s = n - 1.$$

3. Partitioning on $z_{ij} - \hat{\mu} - \hat{\gamma}_j$ (row averages not removed; column averages removed):

$$r = m - 1, \qquad s = n.$$

4. Partitioning on $z_{ij} - \hat{\mu} - \hat{\rho}_i - \hat{\gamma}_j$ (both row averages and column averages removed):

$$r = m - 1, \qquad s = n - 1.$$

5. Partitioning on residuals after removal of row and column averages and p sweeps of the vacuum cleaner:

$$r = m - 1 - p, \qquad s = n - 1 - p.$$

A More Practical Tabulation

For a table of random normal deviates ($\sigma^2 = 1$), we have, according to Eq. (21),

$$\sum_{k=1}^{K} \nu_k = E\left(\sum_{k=1}^{K} \hat{\theta}_k^2\right). \tag{23}$$

But,

$$\sum_{k=1}^{K} \hat{\theta}_k^2 = \sum_i \sum_j d_{ij}^2 = (m-1)(n-1)\frac{\sum_i \sum_j d_{ij}^2}{(m-1)(n-1)}. \tag{24}$$

Hence:

$$\sum_{k=1}^{K} \nu_k = (m-1)(n-1)E\left[\frac{\sum \sum d_{ij}^2}{(m-1)(n-1)}\right]. \tag{25}$$

Since the model is additive and $\sigma^2 = 1$, we have

$$E\left[\frac{\sum \sum d_{ij}^2}{(m-1)(n-1)}\right] = 1. \tag{26}$$

Hence:

$$\sum_{k=1}^{K} \nu_k = (m-1)(n-1). \tag{27}$$

Equation (27) suggests that we compute the values $\nu_k/(m-1)(n-1)$ for each k. The sum of these values must be unity. Alternatively, one can express these quantities as percentages by calculating

$$100\,\frac{\nu_k}{(m-1)(n-1)}, \tag{28}$$

and the sum of these values must be 100. The quantities (28) are tabulated in Table IV. They lend themselves much more readily to accurate interpolation than those of Table II.

Use of Degrees of Freedom in the Analysis of Non-additive Data

Suppose that we have a set of observations z_{ij} tabulated in an $m \times n$ matrix. We can now perform the analysis of variance shown in Table I, using the ν_k values defined above. Let us observe the calculated mean squares. If the real model underlying our data is additive, the mean square for η_{ij} will be an estimate of σ^2, the error-variance of the data. But then, according to our method

of calculation of the ν_k, each term $\hat{\theta}_k^2/\nu_k$ will also be an estimate of σ^2. The value of the mean square for η_{ij}, as well as for every mean square resulting from the partitioning of η_{ij} will have the same order of magnitude.

If, on the other hand, the real model contains a single term $\theta u_i v_j$, then only one of the mean squares of the partitioning will reflect the existence of this term (and will be considerably larger than σ^2) while all other mean squares will only be estimates of σ^2. Since, by the very nature of the calculation of the $\hat{\theta}_k^2$, they always appear in decreasing order (see Appendix A), the large mean square of the partitioned interaction will be the first one. A similar reasoning applies if two or more multiplicative terms occur in the real model.

Thus, by observing the magnitude of the mean squares in the partitioned interaction, a judgment can be made on how far the partitioning process should be carried. At this time, no analog to the F-test has been developed, but it turns out that in most practical situations, the behavior of the mean squares leaves little doubt as to which of them reflect real terms in the model and which can be considered as the result of random fluctuations. This was strikingly illustrated in the numerical example discussed earlier, and has been verified by the author in many practical applications.

Precision of Mean Squares, Degrees of Firmness

An interesting conclusion can be reached from a study of Tables II and III. Consider, for example, the entries for $k = 1$, $r = 9$ and $s = 7$. We find:

$$E(\hat{\theta}^2) = 24.35$$

$$\sigma(\hat{\theta}^2) = 5.30$$

The value $\hat{\theta}^2$ is a sum of squares. For an ordinary sum of squares in analysis of variance, say SS, we have, under the null hypothesis,

$$\mathrm{SS} = \frac{\chi_\varphi^2}{\sigma^2}, \tag{29}$$

where χ_φ^2 is a central chi-square variate with φ degrees of freedom. Hence, for $\sigma^2 = 1$, we have

$$E\ (\mathrm{SS}) = E(\chi_\varphi^2) = \varphi, \quad \text{and} \quad \mathrm{Var}\ (\mathrm{SS}) = \mathrm{Var}\ (\chi_\varphi^2) = 2\varphi.$$

Consequently,

$$\frac{[E(\mathrm{SS})]^2}{\mathrm{Var}\ (\mathrm{SS})} = \frac{\varphi}{2},$$

or

$$\varphi = \frac{2}{(\mathrm{CV}_{\mathrm{SS}})^2}, \tag{30}$$

where $\mathrm{CV}_{\mathrm{SS}}$ is the coefficient of variation, $\sigma_{\mathrm{SS}}/E\ (\mathrm{SS})$, of SS. Thus, φ is not only a divisor to derive a mean square from a sum of squares, but also a measure of the stability, or "firmness", of the sum of squares, since it is inversely proportional to the square of its coefficient of variation.

The concept of "degrees of firmness" and its expression in terms of the quantity $2/(CV_{ss})^2$ are due to J. Tukey [1]. *For any multiple of a chi-square variate, the degrees of firmness are identical with the degrees of freedom.*

TABLE V

Comparison of Degrees of Freedom (D. Fr.) and Degrees of Firmness (D. Fi.)

k	r \ s	3 D.Fr.	3 D.Fi.	9 D.Fr.	9 D.Fi.	19 D.Fr.	19 D.Fi.
1	3	6.7	8.6	16.	22.	28.	41.
	9	16.	22.	28.	52.	45.	87.
	19	28.	41.	45.	87.	66.	140.
	99	120.	172.	153.	343.	192.	496.
2	3	2.0	4.2	8.0	17.	18.	38.
	9	8.0	17.	19.	47.	34.	92.
	19	18.	38.	34.	92.	54.	169.
	99	98.	186.	134.	418.	171.	666.
3	3	.28	1.12	3.5	8.1	11.	23.
	9	3.4	8.1	13.	41.	26.	103.
	19	11.	23.	26.	103.	45.	178.
	99	79.	148.	119.	465.	156.	755.

Let us now calculate the degrees of firmness for our numerical example above.

$$\frac{2}{(CV_{\hat{\theta}2})^2} = \frac{2}{(5.3/24.3)^2} = 42.2.$$

Thus, for $\hat{\theta}^2$, the degrees of firmness are considerably larger than the degrees of freedom, 24.3. Table V lists the degrees of firmness as well as the degrees of freedom for some of the combinations listed in Table II. It is seen that the degrees of firmness are larger in all cases than the degrees of freedom, and in many cases considerably larger. It follows that the mean squares obtained in our partitioning of interaction are generally more "firm" (i.e., have smaller coefficient of variation) than ordinary mean squares with the same number of degrees of freedom.

THE PROPOSED METHOD AS A DIAGNOSTIC TOOL IN DATA ANALYSIS

The model chosen for the illustrative example, (Eq. (E1)), contains two multiplicative terms of the type $\theta u_i v_j$. However, if one is satisfied with a slightly poorer fit, Eq. (E3) is appropriate. The point of interest here is that the method of analysis described in this paper is of sufficient generality to automatically lead to simpler models when these are applicable.

The following is a further illustration of the value of our model as a diagnostic tool.

In a study of the performance of spectrophotometers, Wernimont [8] uses

the method of principal components for the analysis of a table of 6 rows and 20 columns. This method is equivalent to the one discussed in the present paper, except that a) neither row effects nor column effects are eliminated prior to the partition of the data into the sum of multiplicative terms, and b) the degrees of freedom are not properly allocated to the terms in the partitioned sum of squares. Application of our method of analysis to Wernimont's data reveals at once that:

a. A single multiplicative term $\theta u_i v_j$ suffices;
b. $u_i = r_i$;
c. $v_j = c_j$.

This leads directly to the following model (see [3]):

$$z_{ij} = z_0 + \frac{\theta}{RG}[(\mu + \rho_i) - z_0][(\mu + \gamma_j) - z_0] + \epsilon_{ij} ,$$

where z_0 is a constant equal to $\mu - RG/\theta$. Wernimont's conclusion, which he expressed in geometric language, is that the data obey the model $z_{ij} = A_i B_j$, which is consistent with the result of our analysis provided that $z_0 = 0$. (The estimate for z_0 in our analysis is quite small, of the order of 0.001 absorbance units, which is less than the standard deviation of experimental error.) Thus, our method of analysis leads readily to the simple model applicable to these data, without the risk of unwarranted prior assumptions.

Appendix

There exist several algorithms† for obtaining the quantities $\hat{\theta}_k$ and the vectors $[u_{ki}]$ and $[v_{kj}]$. One of these is as follows:

1. Multiply the matrix (d_{ij}) by its transpose $(d)'$. Let $P = (d)(d)'$.
2. The eigenvalues of P are the quantities $\hat{\theta}_1^2$, $\hat{\theta}_2^2$, $\cdots$ $\hat{\theta}_k^2$.
3. The vectors $[u_{ki}]$ are the eigenvectors of P, associated with the eigenvalues $\hat{\theta}_1^2$, $\hat{\theta}_2^2$, $\cdots$
4. The vectors $[v_{kj}]$ are calculated by the formula:

$$v_{kj} = \frac{1}{\hat{\theta}_k} \sum_i d_{ij} u_{ki} .$$

It should be noted that the procedure above is indicated when $m \leq n$. For the case $m > n$, it is more expeditious to compute the matrix $P^* = (d)'(d)$. The non-zero eigenvalues of P^* are the same as those of P. The eigenvectors of P^* are the $[v_{kj}]$, and the vectros $[u_{ki}]$ are then obtained by the formula:

$$u_{ki} = \frac{1}{\hat{\theta}_k} \sum_j d_{ij} v_{kj} .$$

† See for example: Golub, G. H. and Reinsch, C., Singular Value Decomposition and Least Squares Solutions, Technical Report No. CS 133. Computer Science Department, School of Humanities and Sciences, Stanford University (May 1969).

ACKNOWLEDGMENT

The author wishes to thank the referee for very valuable comments and suggestions.

REFERENCES

[1] BLACKBURN, R. B., and TUKEY, J. W. (1958). The Measurement of Power Spectra from the Point of View of Communication Engineering—Part I. *The Bell System Technical Journal 37*, 185–282.

[2] MANDEL, J. and MCCRACKIN, F. L. (1963). Analysis of Families of Curves. *Journal of Research—National Bureau of Standards, Section A, Physics and Chemistry 67*A, 259–267.

[3] MANDEL, J. (1969). A Method for Fitting Empirical Surfaces to Physical or Chemical Data. *Technometrics 11*, 411–429.

[4] MANDEL, J. (1969). The Partitioning of Interaction in Analysis of Variance. *Journal of Research—National Bureau of Standards, Section B, Mathematical Sciences, 73B*, No. 4.

[5] MANDEL, J. (1970). Distribution of Eigenvalues of Covariance Matrices of Residuals in Analysis of Variance. *Journal of Research—National Bureau of Standards, Section B, Mathematical Sciences 74B*, No. 3.

[6] OSBORNE, N. S., MCKELVY, E. C., and BEARCE, H. W. (1913). Density and Thermal Expansion of Ethyl Alcohol and its Mixtures with Water. *Bulletin of the Bureau of Standards, 9.*

[7] TUKEY, J. W. (1962). The Future of Data Analysis. *Ann. Math. Statist. 33*, 1–67.

[8] WERNIMONT, G. (1967). Evaluating Laboratory Performance of Spectrophotometers. *Analytical Chemistry 39*, 554–562.