(NASA-TM-73874) CHAIN POOLING MODELING SELECTION AS DEVELOPED FOR THE STATISTICAL ANALYSIS OF A ROTOR BURST PROTECTION EXPERIMENT (NASA) 60 p HC A04/MF A01 CSCL 12A G3/65 N78-16735 Unclas 02582

NASA TECHNICAL MEMORANDUM

NASA TM-73874

NASA TM-73874

"CHAIN POOLING" MODEL SELECTION AS DEVELOPED FOR THE STATISTICAL ANALYSIS OF A ROTOR BURST PROTECTION EXPERIMENT

by Arthur G. Holms
Lewis Research Center
Cleveland, Ohio 44135

TECHNICAL PAPER presented at the
One hundred and thirty-seventh Annual Meeting
of the American Statistical Association
Chicago, Illinois, August 15-18, 1977

CONTENTS

"CHAIN POOLING" MODEL SELECTION AS DEVELOPED FOR THE STATISTICAL ANALYSIS OF A ROTOR BURST PROTECTION EXPERIMENT

by Arthur G. Holms

Lewis Research Center

SUMMARY

A statistical decision procedure called chain pooling had been developed for model selection in fitting the results of a two-level fixed effects full or fractional factorial experiment not having replication. The basic strategy included the use of one nominal level of significance for a preliminary test and a second nominal level of significance for the final test

The subject has been reexamined from the point of view of using as many as three successive statistical model deletion procedures in fitting the results of a single experiment. The new investigation consisted of random number studies intended to simulate the results of a proposed aircraft turbine engine rotor burst protection experiment. As a conservative approach, population model coefficients were chosen to represent a saturated 2^4 experiment with a distribution of parameter values unfavorable to the decision procedures.

Three model selection strategies were developed, namely, (1) a strategy to be used when the experimenter anticipates a large error variance (coefficients of variation in the neighborhood of 65 percent), (2) a strategy to be used when the experimenter anticipates a small error variance (coefficients of variation of 4 percent or less), and (3) a security regret strategy to be used in the absence of such prior knowledge.

INTRODUCTION

The two-level, fixed effects, full or fractional-factorial design of experiment, without replication, is the appropriate design for those situations where the experiment is very expensive or time consuming. An example of costly experimenting is provided by the destructive testing of simulated aircraft components, as in the rotor burst protection testing described by Mangano (1977). New rotor burst protection investigations are planned to measure the containment efficiencies of composite structures. One such investigation is planned as a two-level fractional-factorial experiment. The description (Holms 1977) of that experiment is illustrative of one area of applicability of the results of the present investigation.

If such an experiment is performed and t observations are obtained from t orthogonal experimental conditions, the appropriate emprical equation for representing the results can have as many as t terms, each with a coefficient that has been fitted to the data. When this is done, a question that should be asked is: "Can the predictive accuracy be improved if some of the terms are deleted?" The fact that some of the terms might degrade the predictive accuracy of a fitted equation was recognized by Walls and Weeks (1969) but they gave no procedure for identifying such terms

A method for the sequential deletion of terms that was intended to reduce the prediction error was given by Kennedy and Bancroft (1971). Their method assumed that the experimenter has a prior established order for subjecting the terms to a sequence of significance tests. Unfortunately, in many experimental situations, there is no basis for establishing a prior order, and in such cases an order statistics procedure is appropriate. An order statistics approach for significance testing was used in a pair of related papers by Daniel (1959) and by Birnbaum (1959). They were not then seeking to minimize prediction errors.

For model selection procedures used with small saturated experiments (experiments designed to have only as many experimental conditions as there are model parameters to be fitted), the analysis should begin with a minimum number of estimable terms being sacrificed to form a denominator for the test statistic. A procedure using m-terms sacrificed, where m can be as small as one, was investigated by Holms and Berrettoni (1969). The procedure was a form of subset regression using backward elimination The object was to delete terms in a manner where some control was maintained over the probabilities of Type 1 or Type 2 decision errors. The procedure was called chain pooling and used a strategy (m_p, α_p, α_f) where m_p was the number of terms initially sacrificed to the test statistic, α_p was the nominal level of a preliminary test of significance where "insignificant" resulted in another term (mean square) pooled into the denominator of the test statistic, and α_f was the nominal level of the final test of significance, for the inclusion of terms in the model.

The minimizing of prediction error was the object of further investigation of the chain pooling strategy (m_p, α_p, α_f) as described by Holms (1974). Whereas that investigation had assumed that only one cycle of analysis would be used, a suggestion given by Holms and Berretoni (1969) was that more than one cycle of analysis should be used. The specific suggestion was that an analysis be performed with the strategy (m_p, α_p, α_f) where $m_p = 1$ to obtain an estimated number of null mean squares, $\hat{n}$, tacitly assuming that the population number of null mean squares is greater than zero. A second analysis would then be performed with a strategy (m_p, α_p, α_f) where m_p is an integer approximation to the product $r_\eta \hat{\eta}$ where r_η had been empirically optimized at about $r_\eta = 0.7$.

The purpose of the present investigation is to use Monte Carlo methods to optimize a combined procedure that might contain more than one analysis cycle, where the procedure is to be optimized for minimum prediction error. Such a procedure would seem to be worthwhile in view of the large number of aerospace research and development programs where:

(1) There are a large number of controlled variables

(2) The experimenting is time consuming or expensive

(3) The dependent variable has an error variance that might be large

Based on the results, some specific model selection procedures are recommended.

PRIOR WORK

An early practice described by Davies (1956, p. 286) consists of pooling some arbitrary number of those mean squares that represent higher order terms into an estimate of error variance. When this is done, some terms of important predictive value might be deleted from the model. Furthermore, such mean squares might unduly inflate the estimate of the error variance, thus reducing the sensitivity of the subsequent procedure, and resulting in the deletion of additional important terms from the model.

The preservation of sensitivity, when pooling mean squares into the estimate of error variance, has been an object of the procedure of Daniel (1959, and of Wilk, Gnanadesikan, and Freeny (1963).

Daniel (1959) uses the absolute values of the effect estimates as order statistics. They are plotted on probability paper and the result is called a half-normal plot. In addition to conditional structuring of the ANOVA model, Daniel's objectives included the determination of "bad values, heteroscedasticity, dependence of variance on mean, and some types of defective randomization, . . .". The half normal plot, combined with a background of experience, might provide a method by which a skillful user could pass judgement on the results of an experiment. Daniel concluded that such a plot can be used to make judgements about the reality of the largest effects observed only if a small proportion of the effect estimates represent real effects. Birnbaum (1959) investigated procedures related to half-normal plotting. His results on ". . . the probabilities of the various possible sorts of errors . . ." are limited to the single largest order statistic. He surmised that if only a small number of effect estimates have nonzero means, the power and sensitivity properties will tend to hold approximately.

The procedure of Wilk, Gnanadesikan, and Freeny (1963) if used with 2^{ℓ} treatments is benefitted if some subjective or prior knowledge is used to decide that η of the $2^{\ell} - 1$ mean squares do not contain real effects, or that $\rho = 2^{\ell} - \eta - 1$ mean squares do contain real effects. As was shown by Wilk et al. (1963), the procedure is not robust against errors in guessing the value of η, and η must be guessed if the prior knowledge is lacking, which is an assumption of the present investigation.

Daniel and Birnbaum have limited their results to experiments where only a small proportion of the effects are anticipated to be significant. On the other hand, situations can exist where the experimenter might use a two-level fractional factorial experiment designed such that a large proportion of the effects are significant. A need remains for a decision procedure that sacrifices only a small number of terms to produce a test statistic, and then uses the test statistic to minimize prediction errors.

Allen (1971) referred to literature that used the residual sum of squares as a criterion for choosing regressors. He pointed out that there are at least two objections to using such a criterion

1. If the residual sum of squares were the only criterion, then all of the regressors would be used and there would be no motivation for subset regression.

2. The residual sum of squares is not directly related to the "natural" loss function which is the mean square error of prediction.

On the other hand, the methods of Allen obtain a criterion called PRESS by predicting each observation from all the other observations. The present investigation is concerned with small saturated experiments where each observation is more or less crucial to the estimation procedure, particularly for predictions at the set of regressor values corresponding to a given observation. The PRESS criterion is therefore believed to be inappropriate to subset regression for small saturated experiments.

MULTISTAGE DECISION PROCEDURE

Assumptions

The assumptions are as follows.

(1) The model to be fitted is linear in the unknown parameters.

(2) The errors of the observations are independently normally distributed random variables with a zero mean and a constant variance.

(3) Orthogonal estimators are available for estimating the unknown parameters of the linear model. (This orthogonality can be the result of the design of the experiment that furnished the observations, or it can

be the result of an orthogonalizing transformation of the terms of the equation, as discussed by Holms (1974).)

(4) An appropriate criterion of the goodness of a subset regression procedure is the smallness of the largest of the prediction error mean squares over the points of the experiment.

(5) There is no replication available for an estimate of the "pure" error variance.

The first four of the five preceding assumptions are funadmental to the rationale of the method; however, data has been given (Holms and Berrettoni (1967)) for believing that the method is robust against the normality part of the second assumption. The fifth assumption merely acknowledges the possibility that an altogether different method might be preferred in the presence of pure replication.

The original investigation of chain pooling had been concerned with three sizes of experiments, namely, experiments furnishing 16, 32, or 64 observations (Holms and Berrettoni (1967)). The simulations had shown drastically reduced decision error probabilities, or the equivalent, greatly improved information efficiencies, for the larger experiments. Such results suggest that the method of analysis is relatively less critical for the larger experiments, and the methods described (Holms and Berrettoni (1967)) are therefore believed to be adequate for producing small prediction errors with experiments providing 32 or more observations.

For relatively saturated experiments that are smaller than 16 observations, the opinion is offered that such experiments are too small to provide both (1) good estimates of model coefficients and (2) a good test statistic, in cases where random errors are large enough to call for a statistical decision procedure. In other words, saturated experiments with less than 16 observations should be fitted with models having a fixed number of terms with no use of conditional modeling.

Consistent with the preceding remarks, the simulations of the present investigation were all performed with experiments containing 16 observations in the belief that such experiments are large enough to justify the use of a statistical decision procedure, but small enough so that the precise optimization of the decision procedure would be quite beneficial.

Strategy

The generalization of the strategy (m_p, α_p, α_f) to be investigated is the strategy (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f), where the symbols are defined in appendix A.

SIMULATION PROCEDURE

Population Model

The experiment is assumed to be a two-level, fixed effects, fractional-factorial experiment, where the independent variables are qualitative or quantitative, and they are controlled to have negligible error. In such situations, the independent variables are often assigned the values of +1 for the "upper" level and -1 for the "lower" level. An example of an appropriate model equation for the population mean value of the response in the case of four independent variables is

$$E(Y) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_{12} x_1 x_2 + \alpha_3 x_3 + \alpha_{13} x_1 x_3 + \alpha_{23} x_2 x_3$$
$$+ \alpha_{123} x_1 x_2 x_3 + \alpha_4 x_4 + \alpha_{14} x_1 x_4 + \alpha_{24} x_2 x_4 + \alpha_{124} x_1 x_2 x_4$$
$$+ \alpha_{34} x_3 x_4 + \alpha_{134} x_1 x_3 x_4 + \alpha_{234} x_2 x_3 x_4 + \alpha_{1234} x_1 x_2 x_3 x_4 \tag{1}$$

The subsequent discussion assumes that an equation such as the preceding equation will be fitted to the results of a two-level experiment where the x's are "design values," namely, the high level of x_k is represented by $x_k = +1$ and the low level of x_k is represented by $x_k = -1$.

The single observation value of the response is assumed to occur according to the model

$$Y = E(Y) + e \tag{2}$$

where e is the independently normally distributed random error, with

$$E(e) = 0; \ V(e) = \sigma^2 \tag{3}$$

The experiment is assumed to be a two-level, fixed effects, fractional-factorial experiment where g is the number of independent variables, and the experiment is assumed to be a 2^{-h} fractional replicate of the full factorial experiment. The observations are assumed to result one-for-one from the treatments, and their number is

$$t = 2^{g-h}$$

The observations are used to compute estimates of parameters such as the parameters of equation (1) and also to compute mean squares, typically by Yates' method (Davies, 1956, p. 263).

Assume that the mean squares Z_i have been computed in Yates' order and in this order are labeled $Z_1, Z_2, \ldots, Z_t$. The coefficients, α's, of equation (1) are in Yates' order. If they are subscripted in that order and then given the symbol β_i, the expectations of the Z_i are

$$\left.\begin{aligned} E(Z_i) &= \sigma^2 + 2^{g-h}\beta_i^2 \\ i &= 1, 2, \ldots, t \end{aligned}\right\} \tag{4}$$

The random variables Z_i/σ^2 are noncentral chi-square variables having one degree of freedom (Kendall, M. G. and Stuart, A., 1961, p. 227). Let λ_i be the noncentrality parameter. It is related to parameters already defined by

$$\lambda_i = \frac{2^{g-h}\beta_i^2}{\sigma^2} \tag{5}$$

A stepwise multiple decision procedure is to be used to delete estimated coefficients from an equation such as the one illustrated by equation (1). The situation for which the operating characteristics of the statistical decision procedure will be optimized is assumed to be as follows:

The t single degree of freedom mean squares are assumed to have been drawn from t populations. An unknown number, ρ, of the populations (aside from the one associated with β_1) have <u>real effects</u> ($\lambda_i > 0$) and the η other populations are <u>null populations</u> (η populations have $\lambda_i = 0$). A number, $\hat{\rho}$, of the populations (other than the one associated with β_1) are to be <u>selected</u> for retention of model coefficients in equations like (1), hoping they have $\lambda_i > 0$.

The relative magnitudes of the nonzero population parameters to be tested can be expressed by dividing squares of the coefficients by the mean value of the sum of the squares, namely, where the quantities δ_k^2 give these relative magnitudes,

$$\left.\begin{aligned} \delta_k^2 &= \frac{\beta_k^2}{\frac{1}{\rho}\sum_{k=1}^{\rho}\beta_k^2} \\ k &= 1,\ 2,\ \ldots,\ \rho \end{aligned}\right\} \tag{6}$$

From equations (6)

$$\frac{1}{\rho}\sum_{k=1}^{\rho}\delta_k^2 = 1 \tag{7}$$

Unfavorable Population Model

The basic chain pooling concept was investigated for the purpose of minimizing prediction error as described by Holms (1974). That investigation was concerned with the fitting of models to fractional factorial experiments under the condition of population functions of irregular shape. The emphasis of the present investigation is on the condition where the relative values of the population coefficients are all unfavorable to the deletion procedure. For reasons given by Holms and Berrettoni (1969) this is done by proportioning the squares of the β_k values to the expected values of the order statistics of a single $\chi^2_{(1)}$ distribution.

Where the ζ_k are the expectations of the $\chi^2_{(1)}$ order statistics in increasing order from a sample of size ρ, the β_k^2 in the order of Yates' computation beyond the first were given a decreasing order by setting

$$\left.\begin{aligned} \delta_k^2 &= \zeta_{\rho-k+1} \\ k &= 1,\ \ldots,\ \rho \end{aligned}\right\} \tag{8}$$

Let λ be the arithmetic mean of the noncentrality parameters.

$$\lambda = \frac{1}{\rho} \sum_{k=1}^{\rho} \lambda_k \tag{9}$$

From equations (5) and (9)

$$\lambda = \frac{1}{\rho} \sum_{k=1}^{\rho} \frac{2^{g-h}\beta_k^2}{\sigma^2} \tag{10}$$

Using the definition of δ_k^2 of equations (6),

$$\beta_k^2 = \delta_k^2 \frac{1}{\rho} \sum_{k=1}^{\rho} \beta_k^2 \tag{11}$$

From equations (10) and (11)

$$\beta_k^2 = \frac{\delta_k^2 \lambda \sigma^2}{2^{g-h}} \tag{12}$$

From equations (8) and (12)

$$\beta_k^2 = \lambda \frac{\sigma^2}{2^{g-h}} \zeta_{\rho-k+1} \tag{13}$$

or

$$\beta_k = \left(\lambda \frac{\sigma^2}{2^{g-h}} \zeta_{\rho-k+1}\right)^{1/2} \tag{14}$$

Expectations of order statistics from a gamma distribution with scale parameter one, shape parameter 1/2 and many sample sizes have been tabulated (Harter, H. L. (1964)). Multiplying such values by 2 gives the expectations of the order statistics of the central $\chi^2_{(1)}$ distribution. Such expectations, for a sample of size ρ, provide the values called for by the definition of the ζ_k.

With the distribution of ζ_k values fixed as just described, the population parameters β_k, as given by equation (14), depend only on the single observation error variance σ^2 and the mean, λ, of the noncentrality parameters.

The simulated experiments were performed and the decision procedures were investigated with η mean squares having $\lambda_i = 0$ and with $\rho = 2^{g-h} - 1 - \eta$ mean squares having the unfavorable set of λ_i values (eq. (5)) that were just described.

Observation Simulation

The number of treatments is

$$t = 2^{g-h} \tag{15}$$

In accordance with model equations such as (1) the t observations resulting from t treatments are given by equation (2) with $E(Y_k) = \mu_k$ as the variates

$$\left.\begin{aligned} y_1 &= \mu_1 + e_1 \\ y_2 &= \mu_2 + e_2 \\ &\cdots\cdots\cdot \\ y_t &= \mu_t + e_t \end{aligned}\right\} \tag{16}$$

In the analysis of an experiment, the application of Yates' method estimates each of the t parameters that occur in an equation such as (1).

Furthermore, from equations (12), the nonzero values of the β_i are

$$\beta_k = \left(\frac{\lambda\sigma^2}{2^{g-h}}\right)^{1/2} \delta_k \tag{17}$$

where

$$k = 1,\ 2,\ 3,\ .\quad .,\ \rho$$

$$i = 1,\ 2,\ .\ .\ .,\ t$$

and

$$i = k + 1 \qquad \text{for } k = 1,\ .\quad ,\ \rho \tag{18}$$

and $\beta_i = 0$ otherwise Let

$$\theta^2 = \frac{\lambda}{2^{g-h}} \tag{19}$$

From equations (17) and (19)

$$\beta_k = \theta\sigma\delta_k \tag{20}$$

Any particular strategy (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f) will be evaluated for an array of populations having m_ρ unique values of ρ, namely, ρ_m, $m = 1, 2, \ldots, m_\rho$ and ℓ_λ unique values of the mean noncentrality parameter, λ, namely, λ_ℓ, $\ell = 1, 2, \ldots, \ell_\lambda$.

Because of the freedom to choose the values λ_ℓ and because λ_ℓ is a scale parameter on σ^2 (eq. (17)) an investigation of the effect of variations in σ^2 is superfluous, and σ^2 will be set equal to one.

As implied by equation (16) the computer program must perform operations equivalent to the following:

$$y_1 = \mu_1 + e_1 \tag{21}$$

$$y_2 = \mu_2 + e_2$$

$$\cdot \quad \cdot \cdot \quad \cdot$$

$$y_{16} = \mu_{16} + e_{16} \tag{22}$$

As previously stated, the $\hat{b}_i$ are to be computed from the y_i by Yates' method. With the $\hat{b}_i$ computed, the reversed Yates method of Duckworth (1965) is used to compute predicted values of Y, namely, $\hat{y}_i$ The prediction errors are then

$$e_{pi} = \hat{y}_i - \mu_i \tag{23}$$

<u>Conditions Investigated</u>

As developed in the subsection on strategy in the main section on MULTISTAGE PROCEDURE, the statistician's strategy consists of the components (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f).

As developed in the subsection labeled Unfavorable Population Model, nature's strategy consists of the number, ρ, of non-null population parameters, and the mean noncentrality parameter, λ. As given by equation (19) the relation between the mean noncentrality parameter, λ, and the scale parameter θ is

$$\theta^2 = \frac{\lambda}{2^{g-h}}$$

For the experiment with 16 treatments, $2^{g-h} = 16$ and the related values of θ and λ investigated are as given in table I.

In general, the smaller the number of null mean squares, η, the greater will be the probability of decision errors. This was illustrated in figure 4 of the paper by Holms and Berrettoni (1969). Thus the most difficult situation arises for $\eta = 0$. For the 2^{g-h} experiment with $g - h = 4$, and where the β_1 term (zero degree term) is not subjected to testing, the condition equivalent to $\eta = 0$ is the condition $\rho = 15$.

A real-life example of a $2^{g-h} = 2^{5-1}$ experiment with $\rho = 15$ was presented by Daniel (1959) in figure 19 of his paper. As stated by Daniel concerning such a situation:

> "The 15 effects appear to fall into a nearly normal distribution with their own variance, much larger than the error variance. The reality of many of these terms is proved by later experimental work. There may be some sort of central limit theorem analog operating here on the (discrete) population of effects, but no success can be reported in its formulation."

One interpretation of Daniel's discussion is that if an experimenter designs his experiment and attempts to select the levels of the controlled variables so that measurable effects will be achieved, then a preponderance of the fitted coefficients of the model equation will have absolute magnitudes fairly close to some central value, and a very small number of coefficients will have much larger and much smaller values, and still other coefficients will be scattered between these two extremes. In other words, an approximation to the normal distribution should be expected for the coefficients of a model equation resulting from experiments where controlled variables were set with some degree of prior knowledge of their effects.

In the light of the preceding considerations, a nature's strategy with $\rho = 15$ and a normal distribution of model parameters would seem to be a highly likely strategy, and correspondingly, a statistician's strategy optimized against such a nature's strategy should be thought of as an <u>approximate Bayes strategy</u>. As developed by Holms and Berrettoni (1969), such a normal distribution of model parameters is represented by the parameter distributions of table II and these distributions are highly unfavorable to the statistical decision procedure. A procedure optimized against $\rho = 15$ and the distribution of ξ_j of table II may therefore also be regarded as an <u>approximate security strategy</u>. Such a nature's strategy (table II) will therefore be chosen as the strategy against which the statistician's strategy will be optimized.

Consider the fitting of a model equation such as that illustrated by equation (1) to the results of a two-level, fixed effects, full or fractional factorial experiment. The independent variables can be standardized so that all the x-values are either +1 or -1. With t treatments and t observations, the estimators of the coefficients are all of the form

$$\hat{b}_i = \frac{a_1 y_1 + a_2 y_2 + \dots + a_t y_t}{t} \tag{24}$$

where the $a_1, a_2, \dots, a_t$ all have values of +1 or -1 The single observation error variance of the Y-values is assumed to be σ^2 and because the $\hat{b}_i$ are formed from a linear combination of t independently distributed Y-values (eq. (24)), the variance of any $\hat{b}_i$ is

$$V(\hat{b}_i) = t\,\frac{\sigma^2}{t^2} = \frac{\sigma^2}{t} \tag{25}$$

With equation (1) written as a prediction equation, the predicted value of Y, at any one of the points of the experiment, is a linear combination of t independently distributed $\hat{b}_i$ values, each (eq (25)) with variance σ^2/t. With appropriate change of notation, the linear combination implied by equation (1) is written

$$\hat{Y}_i = \sum_{j=1}^{t} c_j \hat{b}_j \tag{26}$$

and from the definition of the x-values in equation (1), the c_j-values in equation (26) are either +1 or -1. Because the $\hat{b}_j$ are independently distributed, and because the c_j are all +1 or -1, the variance of $\hat{Y}_i$ is

$$V(\hat{Y}_i) = \sum_{j=1}^{t} V(b_j) = t\,\frac{\sigma^2}{t} = \sigma^2 \tag{27}$$

From equation (27), the reduction of $V(\hat{Y}_i)$ achievable by deleting terms is σ^2/t for each term deleted.

On the other hand, if equation (1) is the population model, and if the x-values are all +1, the bias in $\hat{Y}$ is increased by the amount of β_i for each β_i value that is deleted. Thus an optimal strategy to minimize the squared error of $\hat{Y}$ should not only delete all terms for which the population β_i is zero, it should also delete at least all terms for which the bias contribution to mean square error is less than the variance contribution.

As indicated by equation (1), for that point of the experiment where all of the x-values have the value +1, the expected value of Y_i would take on its greatest absolute value, which would be

$$E(Y_1)_{max} = \sum_{j=1}^{\rho} \theta \delta_j$$

The values of $\sum_{j=1}^{\rho} \theta \delta_j$ for $\rho = 15$ have been listed in table III. Because $\sigma = 1$, these values are also the values of $E(Y_1)_{max}/\sigma$.

Reciprocals of $E(Y_1)_{max}/\sigma$ are here defined as coefficients of variation for the maximum population mean values. From table III such coefficients vary from a high of 64.3 percent (at $\theta = 0.125$) to a low of 4 0 percent (at $\theta = 2\ 000$). This range of such a coefficient of variation suggests that the range of $0.125 \leqq \theta \leqq 2.000$ is an adequately wide range of θ to represent the situations that an experimenter might encounter

EVALUATION CRITERIA

Following the selection of terms (where some of the coefficient estimates are set equal to zero), the predicted values of the dependent variable can be efficiently computed for all of the combinations of the independent values, by the reversed Yates method of Duckworth (1965).

Where e_{oin} are the "observation" errors, namely, the pseudo normal random numbers generated in the n^{th} simulation, the "observations" are given consistently with equations (16) by

$$y_{oi\ell mn} = \mu_{i\ell m} + e_{o\,in} \tag{28}$$

$\ell = 1, 2, \ldots, \ell_\theta$, $m = 1, 2, \ldots, m_\rho$, $i = 1, 2, \ldots, t$.

After the model has been fitted and insignificant terms deleted, the difference between predicted values, $y_{pi\ell mn}$, of the dependent variable for the n^{th} simulation and the population mean will be called the prediction error, and thus it is (consistent with eq. (23))

$$e_{pi\ell mn} = y_{pi\ell mn} - \mu_{i\ell m} \tag{29}$$

$\ell = 1, 2, \ldots, \ell_\theta$, $m = 1, 2, \ldots, m_\rho$, $i = 1, 2, \ldots, t$.

Over the n_e simulations, the sample mean square error of prediction for a given treatment is

$$\bar{e}^2_{pi\ell m} = \frac{1}{n_e} \sum_{n=1}^{n_e} e^2_{pi\ell mn} \tag{30}$$

The maximum of such errors over the treatments is

$$\bar{e}^2_{max} = \bar{e}^2_{p\ell m, max} = \max_{i=1,\ldots,t} \left(\bar{e}^2_{pi\ell m}\right) \tag{31}$$

The mean of the squared error over the simulations and over the points of the space of the experiment is

$$\bar{e}^2_{p\ell m} = \frac{1}{t} \sum_{i=1}^{t} \bar{e}^2_{pi\ell m} \tag{32}$$

Equations (31) and (32) provide two criteria for measuring the effectiveness of a strategy. The particular set of values of strategy parameters that minimizes $\bar{e}^2_{p\ell m, max}$ (as given by eq. (31)) can be called a security strategy, and if the points of the space of the experiment are assumed to be equally likely of being of interest, the particular set that minimizes $\bar{e}^2_{p\ell m}$ can be called an approximate Bayes strategy. For either criterion, the absolute values of squared errors would have been the prime consideration.

An example of a situation where such definitions of error would be appropriate occurs if the experimentor seeks to maximize some predicted response, such as the strength of a structure as a function of its geometric variables. For such an example, the region of the space of the experiment of greatest interest would be the region in the vicinity of the maximum point, where the function would most likely have its sharpest curvatures and largest errors due to lack of fit. For such an example, the appropriate criterion to be minimized for the choice of a strategy would seem to be the quantity $\bar{e}^2_{p\ell m,\max}$ of equation (31).

The criteria of equations (31) and (32) were evaluated using computer simulations using (in most cases) 1000 experiments. Thus the long run mean squared error of the decision procedures was evaluated. This leaves open the question of how badly a decision procedure might perform in individual cases. One approach to this question is to evaluate the stability of the mean squared errors observed in the simulations. Thus in addition to the criteria of equations (31) and (32) two other criteria for the effectiveness of a strategy were investigated. They are concerned with the stability of the quantities defined by equations (31) and (32). The instability of these criteria can be measured by the variance of the square of the prediction error. If Y is a random variable, the unbiased estimate of the variance of Y from a sample of size n_e is given by

$$\hat{V}(Y) = \frac{1}{n_e - 1}\left[\sum_{n=1}^{n_e} Y_n^2 - \frac{1}{n_e}\left(\sum_{n=1}^{n_e} Y_n\right)^2\right] \tag{33}$$

The random variate of interest is the squared error of prediction, namely, $e^2_{pi\ell mn}$. From equation (33) the estimate of the variance of $e^2_{pi\ell mn}$ is

$$\hat{V}\left(e^2_{pi\ell m}\right) = \frac{1}{n_e - 1}\left[\sum_{n=1}^{n_e}\left(e^2_{pi\ell mn}\right)^2 - \frac{1}{n_e}\left(\sum_{n=1}^{n_e} e^2_{pi\ell mn}\right)^2\right] \tag{34}$$

Equation (34) gives an unbiased estimate of the variance of the squared error over n_e simulations. The maximum of this quantity over the space of the simulated experiments is defined by

$$V\left(e^2\right)_{\max} = \hat{V}\left(e^2_{\ell m}\right)_{\max} = \max_{i=1,\ldots,t}\left[\hat{V}\left(e^2_{pi\ell m}\right)\right] \tag{35}$$

The arithmetic mean of the variance of the squared error over the space of the experiments is defined by

$$\hat{V}\left(\bar{e}^2_{\ell m}\right) = \frac{1}{t} \sum_{i=1}^{t} \hat{V}\left(e^2_{pi\ell m}\right) \tag{36}$$

Two features not present in the computer program POOLES (Holms (1974)) were added to the present program (POOL6U). One of them computes the average number of terms, $\bar{p}_{\ell m}$, selected by the strategy (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f) for each of the values of θ_ℓ, $\ell = 1, \ldots, \ell_\theta$ and for each of the population values of ρ_m, $m = 1, \ldots, m_p$. The other feature computes the ratio of the maximum prediction error to the scale parameter θ. The ratio is computed from θ_ℓ and from the $\bar{e}_{max}$ of equation (31)

$$C_{ee,mx}(\theta, \rho) = \frac{\bar{e}_{p\ell m, max}}{\theta_\ell} \tag{37}$$

COMPUTER PROGRAM

Outline of Program

Computations were performed according to the FORTRAN-4 program POOL6U as listed in appendix B. The antecedents of the program were POOL3U (Holms (1966)), POOLMS (Amling and Holms (1973)), and POOLES (Holms (1974)). The program is outlined and the parts of POOL6U from the earlier programs are identified by the section numbers and titles of appendix B in the table that follows. The table shows that only minor portions of POOL6U were not previously described. The detailed description of POOL6U is mainly limited to sections that are new. Illustrative output is given in appendix C.

Section number	Section title	Prior program
1A	DECLARATIONS AND TABLES	POOLMS
1B	INPUTS AND CONSTANTS	POOLES
1C	POPULATION MEANS	(new)
1D	STRATEGY	(new)
2	SIMULATIONS AND MODEL FITTING	POOL3U

Section number	Section title	Prior program
3	CONSTRUCTION AND ORDERING OF MEAN SQUARES	POOL3U
4	DELETION OF TERMS	(new)
5	PREDICTIONS	POOLES
6	ACCUMULATION OF ERRORS	POOLES
7	DETERMINATION OF MAXIMUM AND MEAN SQUARED ERRORS	POOLES
8	OUTPUT	(new)
9	YATES' METHOD SUBROUTINE	POOLES

Details of Program

Section 1A. - Declarations and tables. - The input of the critical point values (the statistical tables) is the same as in POOLMS (Amling and Holms (1973)). The declarations are similar to POOLES (Holms (1974)). The NAMELIST output was incorporated to facilitate checking of the program and was not otherwise used.

Section 1B. - Inputs and constants. - The constants defining the populations, the experiments, and the sequential deletion strategy, are read from data cards in the following order, with the order of the fields being the same as the order of the symbols in the following description.

Format	Description
(13A6, A2)	REMARK (I), arbitrary literal information such as particular use of program, date of last change, and so forth.
(4I5)	LGMH, NE, MRHO, KODE
(I8, 6F8.3)	LTH, (THETA(L), L=1, LTH)
I4/	KTEM
(10F8.5)	(DELTA(K, M), K=1, KTEM)

Format	Description
	There are as many (10F8.5) cards as are necessary to read (DELTA(K, M), K=1, KTEM). Furthermore, the card sequence for KTEM and DELTA (K, M) is repeated as many times as necessary to satisfy the statement (appendix B) "DO 5 M=1, MRHO."
(4I5, 2F5.3)	MP, KP(1), KP(2), KF, RETA(1), RETA(2) (The associated READ statement is actually in Section 1D.)

Section 1C. - Population means. - After the initial constants have been read, the next major operation is the formation of the population mean values. The number of population sets to be examined during the investigation of a strategy is the number resulting from all combinations of the number, m_ρ, of ρ-values, and the number, ℓ_θ, of θ-values.

With respect to equation (1) all the population model parameters are first set equal to zero with the DO-loop ending at Statement 10. The non-zero values of the β_1 are initially set equal to $\delta_{1,m}$ using the DO-loop ending at Statement 20. The DO-loop ending at statement 20 serves the purpose of equation (20) with $\sigma = 1$ and $\theta = 1$. The value of $\sigma = 1$ is retained, but the adjustment for θ is made after the population mean values have been computed.

With the population β-values (aside from θ) established at statement 20, the objective is to compute the population mean values from the β-values by the reversed Yates' method (Duckworth (1965)). The first step is to reverse the order of the β-values, which is completed at statement 22. The use of the reversed Yates' method then yields the array YMUM (I, M) as completed at statement 40 for all the values of ρ. YMUM (I, M) is therefore an array of population means $\mu_{i,m}$. This array of population means is to be expanded over the mean noncentrality parameters, λ_ℓ, to give the effect of equation (20). This effect is produced on the population mean values by the multiplication

$$\mu_{i,\ell,m} = \mu_{i,m} * \theta_\ell$$

and this operation is completed with the creation of the array YMU (I, L, M) at statement number 49.

The values of YMU are thus indexed over treatments i, $i = 1, \ldots, t$ over arbitrary values of θ_ℓ; $\ell = 1, \ldots, \ell_\theta$: and over the different values of ρ_m; $m = 1, \ldots, m_\rho$.

Section 1D. - Strategy. - In terms of mathematical symbols previously defined, the strategy parameters are functions of numbers that are read at statement 50 as follows:

Argument FORTRAN symbol	Function mathematical symbol
MP	m_p
KP(1)	α_{p1}
RETA(1)	r_1
KP(2)	α_{p2}
RETA(2)	r_2
KF	α_f

For each strategy investigated, the contents of the error arrays are first set equal to zero with the DO-loop ending at statement 99.

Section 2. - Simulations and model fitting. - This section is essentially the same as that described for POOLES by Holms (1974), except that in POOL6U the experiments are simulated and fitted as if they are full-factorial experiments.

Section 3. - Construction and ordering of mean squares. - This section is essentially the same as that described for POOLES by Holms (1974).

Section 4. - Deletion of terms. - This section contains the major changes between POOL6U and POOLES. The major distinction is the provision for more than one iteration of sequential deletion, the maximum number of iterations being three. The new strategy parameters are r_1,

r_2, and α_{p2}. The flow chart for the iterations is shown by figure 1.

The statements from "DO 415 J=1,MPX" to "420 JETA = J-1" are essentially the same as those in POOLMS (Amling and Holms (1973)) from "DO 15 J=1,M" to "20 JETA = JA-1."

The statements beyond 420, and up to and including statement 424 provide for the revision of the number of mean squares initially pooled in the given sequential deletion strategy in accordance with the value assigned to r_1 or r_2.

The statements beyond 424 and terminating with statement 429 set the insignificant model parameters equal to zero, as was done in statements 416 through 419 of POOLES (Holms (1974)).

Section 5. - Predictions. - This section is an abridgment of that described for POOLES (Holms (1974)).

Section 6. - Accumulation of errors. - This section is the same as that of POOLES (Holms (1974)).

Section 7. - Determination of maximum and mean squared errors. - This section is the same as that of POOLES (Holms (1974)) except for the computation of the arithmetic mean, $\overline{\rho}$, of the number of terms (beyond the zero degree term) that are retained by the decision procedure, and the computation of COERMX.

Section 8. - Output. - The output is illustrated in appendix C. The NAMELIST output was incorporated only for program checking.

Section 9. - Yates' method subroutine. - This subroutine is essentially that of part of the main program of POOLMS (Amling and Holms (1973)) except with the last few statements modified so that the subroutine can be used for the direct Yates' method and also for the reversed Yates' method; as was also done in POOLES (Holms (1974)).

The program POOL6U contains some statements from POOLES that are not essential to the present purposes of POOL6U.

SIMULATION RESULTS

Monte Carlo Sample Size

Results of a preliminary investigation of the effect of Monte Carlo sample size on the stability of the empirical results are shown by figure 2. In general, the results converge to a constant when the number of sampled experiments is 1000 or more. Variability of results occurs as the number of sampled experiments is reduced below 1000. All of the strategy comparisons were performed for 1000 sampled experiments.

Large Coefficient of Variation

One of the conclusions of Holms (1974) was that a widely useful three parameter strategy of sequential deletion for minimizing prediction errors consists of the strategy $(m_p, \alpha_p, \alpha_f) = (1, 0.75, 0.10)$. The possibility of performing two analyses in sequence had been suggested by Holms and Berrettoni (1969). That suggestion had included performing an initial analysis with the strategy $(m_p, \alpha_p, \alpha_f)$ with $m_p = 1$. The values of α_p and α_f would be selected according to some attempt to control the probabilities of type 1 or type 2 errors. The initial analysis would yield an estimate of η. Then based on table 3 of Holms and Berrettoni (1969), a second analysis would be performed with m_p approximately equal to 0.67 or 0.75 times the $\hat{\eta}$ from the first analysis. The joint implications of the preceding two sets of results are that for the present investigation with six parameters, a useful strategy might be as follows. $(m_p, \alpha_{p1}, r_1, \alpha_{p_2}, r_2, \alpha_f) = (1, 0.75, 0.700, 0.75, 0.700, 0.10)$. Some Monte Carlo experimenting was started with strategy parameters in the immediate vicinity of the set just listed.

The most adverse value of ρ among nature's strategies is the value $\rho = 15$ Using this value of ρ, the simulations were performed for all the listed values of θ of table I.

The strategies were most often compared in terms of the maximum coefficient of error, $C_{ee,mx}$ defined at equation (37). The largest values of $C_{ee,mx}$ occurred at $\theta = 1.125$. The better strategies for $\theta = 0.125$ are listed in table IV which also lists the values of $C_{ee,mx}$ for both $\theta = 0.125$ and $\theta = 2.000$.

The best strategies of table IV for $C_{ee,mx}$ (0.125) for each value of m_p are listed in table V together with additional information on the operating characteristics of these strategies. These strategies are all good strategies at $\theta = 0.125$, which represents the situation where the model parameters are relatively small (table III) in comparison with the error standard deviation ($\sigma = 1.0$). (From previous discussion in the section on simulations and from table III, the ratio of the maximum population value of the dependent variable to the standard deviation at $\theta = 0.125$ is 1.556 whereas that ratio at $\theta = 2.000$ is 24.898.)

Small Coefficient of Variation

If the statistician's loss function is the maximum relative error over the space of his experiment, $C_{ee,mx}$, then the strategy that is optimal for $\theta = 0.125$ is a security strategy because within the present investigation $C_{ee,mx}$ is larger for $\theta = 0.125$ than for any other value (table I) of θ investigated.

On the other hand, if the statistician's loss function is simply the absolute value of the maximum squared error over the space of his experiment, namely, $\bar{e}^2_{max}$ as defined by equation (31) then that quantity is a maximum within the present investigation at $\theta = 2.000$ and thus the security strategy for such a loss function would be the strategy that minimizes $C_{ee,max}$ (2.000). The strategy that minimizes $C_{ee,max}$ (2.000) is the strategy with no deletion which is symbolized as (m_p, α_{p1}, $r_{\eta 1}$, α_{p2}, $r_{\eta 2}$, α_f) = (0, ____, ____, ____, ____, ____). Some other strategies that were good with $m_p = 1$, $\alpha_{p1} = 1.00$, and $\alpha_f = 1.00$ are listed with resulting values of $C_{ee,mx}$ (0.125) and $C_{ee,mx}$ (2.000) in table VI.

Admissible Strategies

For the purposes of the present investigation, a strategy will be classed either as admissible or as dominated according to its values of $C_{ee,mx}(\theta)$ at both $\theta = 0.125$ and $\theta = 2.000$. A strategy will be said to be dominated if for $\theta = 0.125$ there is another strategy with the same or lesser $C_{ee,mx}$ (0.125) and with a lesser $C_{ee,mx}$ (2.000). A strategy will also be said to be dominated if there is another strategy with the same or lesser $C_{ee,mx}$ (2.000) and with a lesser $C_{ee,mx}$ (0.125).

Any strategy that is not dominated is defined as being admissible For example, within table V, the strategies for $m_p = 1$, $m_p = 3$, and $m_p = 5$ are admissible, whereas the other three strategies are dominated

In addition to the strategies listed in tables IV and VI for $\alpha_{p1} = 1.00$ and $\alpha_f = 1.00$, many strategies with smaller values of α_{p1} and α_f were investigated. Among all of the strategies investigated, the admissible strategies, together with some of their operating characteristics are listed in table VII.

Security Regret Strategy

The strategies of table VII include the strategy $m_p = 0$, which resulted in the smallest observed value of $C_{ee,mx}$ (2.0), namely, 0.5282, and they include the strategy (5, 1.00, ____, 0.05, 0.675, 1.00) which gave the smallest observed value of $C_{ee,mx}$ (0.125), namely, 7.689.

The regret function of a statistical decision procedure, as a function of a parameter θ, is here defined as the excess loss occurring with the procedure at a particular value of θ as compared with the loss that would have occurred had the best statistical decision procedure been used for that particular value of θ. For the purposes of the present investigation a regret function $R(\theta)$ is defined for $\theta = 0.125$ as being the $C_{ee,mx}$ (0.125) for any strategy divided by the value of $C_{ee,mx}$ for the best strategy for that value of θ, namely, 7.689, and $R(\theta)$ is defined for $\theta = 2.000$ as being the $C_{ee,mx}$ (2.000) for any strategy divided by the value of $C_{ee,mx}$ for the best strategy for that value of θ, namely, 0.5282.

Thus,

$$R(0.125) = C_{ee,mx}(0.125)/7.689$$

and

$$R(2.000) = C_{ee,mx}(2.000)/0.5282$$

The single strategy that has the smallest regret function over both $\theta = 0.125$ and $\theta = 2.0$ is defined as the security regret strategy. The security regret strategy is thus the sequential deletion procedure which produces the least increase in prediction error for $\rho = 15$ and an unfavorable distribution of parameters over that prediction error which could have been achieved if the best strategy had been chosen for the given (unknown) value of error variance, σ^2.

Examination of the $R(\theta)$ values of table VII shows that the parameters of the security regret strategy (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f) are (1, 1.00, ____, 0.75, 0.360, 1.00) for which the values of $R(\theta)$ are 1.0966 for $\theta = 0.125$ and 1.0682 for $\theta = 2.000$.

Selection of a Strategy

In summary, if the experimenter wishes to minimize the maximum prediction error over the points of the experiment when the variance error is relatively large ($\theta = 0.125$) the statistician's strategy should be (5, 1.00, ____, 0.05, 0.675, 1.00), which minimizes $C_{ee,mx}$ (0.125), (tables V and VII). If the experimenter wishes to minimize the maximum prediction error over the points of the experiment when the variance error is relatively small ($\theta = 2.000$), the satistician should use the strategy $m_p = 0$, which is to say that no model deletion will be performed, which minimizes $C_{ee,mx}$ (2.000) (table VII).

If the experimenter has no basis for a choice between the two preceding extreme strategies, a security regret strategy with (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f) = (1, 1.00, ____, 0.75, 0.360, 1.00) should be used, in which case the values of the regret function (from table VII) will be $R(\theta) = 1.0966$ if $\theta = 0.125$ and $R(\theta) = 1.0682$ if $\theta = 2.000$. These values of the regret function show that the relative prediction error standard deviation will be increased by at most 9.7 percent over what it would have been if the worst value of θ had occurred and the best strategy against it had been used. Thus the security regret strategy of (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f) = (1, 1.00, ____, 0.75, 0.360, 1.00) must be concluded to be a widely useful strategy.

Variance of Predicted Squared Error

The three strategy selections described in the preceding section are based on a Monte Carlo investigation that reported mean values of prediction errors over 1000 simulations. The quoted results thus tell what the mean long run results will be as a function of strategy selection. The subject of short run results was not discussed. Some insight into the short run performance can be gained by examining the observed values of $V(e^2)_{mx}$. This quantity gives the observed variance, for samples of size 1000, of the maximum squared prediction errors over the simulations, as defined by equation (35). If this variance is relatively small, then operating characteristics such as $C_{ee}(\theta)_{mx}$ are relatively constant from simulation to simulation. But if $V(e^2)_{mx}$ is relatively large, then the short run performance of a strategy could be erratic.

In the case of large coefficients of variation (small values of θ) the strategy performance was not erratic - the values of $V(e^2)_{mx}$ were small for all the strategies of table VII for $\theta = 0.125$. The strategy performance can be erratic for small coefficients of variation (large values of θ). Thus the values of $V(e^2)_{mx}$ were large or small for $\theta = 2.000$, depending on the strategy (table VII). This response to θ shows that the bias component is the component of the prediction error that can be erratic.

The response of $V(e^2)_{mx}$ to the strategy parameters is to be examined, and this will be done for $\theta = 2.000$. First of all, the parameters α_{p1} and r_1 are relatively unimportant. Namely, examination of table VII shows that identical operating characteristics are often obtained despite wide variations of α_{p1} and r_1. Furthermore, the three specific strategy recommendations of the preceding section all use $\alpha_{p1} = 1.00$. For these two reasons, the influence of strategy parameters on $V(e^2)_{mx}$ was examined only for cases with α_{p1} constant at $\alpha_{p1} = 1.00$. Thus the only strategies of table VII that were examined for their influence on $V(e^2)_{mx}$ were those that initiated the first analysis cycle as specified by m_p and α_{p2} and then did or did not perform a second analysis cycle as specified by r_2 and α_f.

Using results from table VII with $\theta = 2.000$ and $\alpha_{p1} = 1.00$, an arbitrary second degree polynomial was fitted. The form of the polynomial was:

$$\ln\left[V(e^2)_{mx}\right] = \beta_0 + \beta_1 m_p + \beta_2 \alpha_{p2} + \beta_3 r_2 + \beta_4 \alpha_f + \beta_5 m_p^2$$
$$+ \beta_6 \alpha_{p2}^2 + \beta_7 r_2^2 + \beta_8 \alpha_f^2 + \beta_9 m_p \alpha_{p2} + \beta_{10} m_p r_2 + \beta_{11} \alpha_{p2} r_2 + \beta_{12} m_p \alpha_f$$
$$+ \beta_{13} \alpha_{p2} \alpha_f + \beta_{14} r_2 \alpha_f \tag{38}$$

The fitting used the backward deletion procedure of Sidik (1972) with a nominal significance level of 0.01. Results rounded to two significant figures are as follows:

$$\ln\left[V(e^2)_{mx}\right] = 9.1 - 0.44\, m_p + 5.5\, \alpha_{p2}$$
$$- 6.7\, \alpha_f - 11.2\, \alpha_{p2}^2 + 9.1\, r_2^2 + 1.0\, m_p \alpha_{p2} \tag{39}$$

The coefficients of equation (39) show that the variability of $C_{ee}(2.000)_{mx}$ decreases with increasing m_p, decreases rapidly with increasing α_{p2}, increases rapidly with r_2, increases with the product of m_p and α_{p2} and decreases with increasing α_f. The range of m_p included the values $m_p = 1, 2, 3, 4, 5$. Thus even for $m_p = 5$ the cross product term in $m_p \alpha_{p2}$ together with the first degree term in α_{p2} are of slightly less influence than the second degree term in α_{p2}.

The values of the strategy parameters all serve to control the number of terms deleted, which fixes the number, $\hat{\rho}$, of terms retained. Thus a given value of $\hat{\rho}$ can result from many different combinations of values of the strategy parameters. From the results exhibited by equation (39) the conclusions as to what combination of strategy values would result in a given $\hat{\rho}$ while minimizing $V(e^2)_{mx}$ are. m_p is unimportant, α_{p2} should be small; r_2 should be small; and α_f should be large.

Thus to minimize both $C_{ee}(\theta)_{mx}$ and $V(e^2)_{mx}$, a strategy should use $\alpha_f = 1.00$ and the smallest values of α_{p2} and r_2 that give acceptable values for $C_{ee}(\theta)_{mx}$ over appropriate values of θ. The choice of m_p is not critical, and the use of $\alpha_{p1} = 1.00$ is generally acceptable.

CONCLUSIONS

An investigation was conducted to determine what statistical techniques should be used for model fitting to the results of a two-level, fixed-effects full- or fractional-factorial, orthogonal experiment with 16 treatments. Multiple sequential deletion strategies involving as many as two preliminary tests and one final test were evaluated, using Monte Carlo techniques, under the criterion of minimum prediction error.

Three strategies were identified as being appropriate depending on the extent of the experimenter's prior knowledge.

1. If the experimenter has prior knowledge that the relative error is relatively small (coefficients of variation of 4 percent or less) then no deletion should be used. The strategy is $(m_p, \alpha_{p1}, r_1, \alpha_{p2}, r_2, \alpha_f) =$ (0, ____, ____, ____, ____, ____,).

2. If the experimenter has prior knowledge that the relative error is quite large (coefficients of variation in the neighborhood of 65 percent) the strategy should immediately delete the five smallest absolute value terms and then test with continued pooling at a nominal test level of 0.05 to estimate a number $\hat{\eta}$ of insignificant terms. The number of terms deleted from the model should then be the integer value of $1 + 0.675\,\hat{\eta}$. The strategy is $(m_p, \alpha_{p1}, r_1, \alpha_{p2}, r_2, \alpha_f) =$ (5, 1.00, ____, 0.05, 0.675, 1.00).

3. Without prior knowledge about coefficients of variation, a security regret strategy should be used. It consists of beginning the sequential deletion with the single smallest absolute value term sacrificed to the test statistic. Testing is done with continued pooling at a nominal test level of 0.75 to estimate a number $\hat{\eta}$ of insignificant terms. The number of terms deleted from the model should then be the integer value of $1 + 0.36\,\hat{\eta}$. The strategy is $(m_p, \alpha_{p1}, r_1, \alpha_{p2}, r_2, \alpha_f) =$ (1, 100, ____, 0.75, 0.36, 1.00).

APPENDIX A

SYMBOLS

Mathematical symbol	FORTRAN name	Description
$\hat{b}_i$	B(I)	estimate of β_i
$C_{ee,mx}$	COERMX	ratio of maximum prediction error to scale parameter, eq. (37)
E (. . .)		expectation of . . .
e	RN(I)	single observation random error
$\bar{e}^2_{max}$	ERSQMX	maximum over treatments of mean square prediction error over simulations
g		number of independent variables
g-h	LGMH	experiment contains 2^{g-h} treatments
h		experiment contains $(1/2)^h$ times number of treatments in full factorial experiment
i, j, k, ℓ, m, n	I, J, K, L, M, N	subscripts
	KF	index number for α_f, Amling and Holms (1973)
	KODE	amount of NAMELIST output desired
	KP	index number for α_p, Amling and Holms (1973)
ℓ_θ	LTH	number of θ values investigated in any computer run
m_p	MP	number of mean squares pooled before testing begins

Mathematical symbol	FORTRAN name	Description
m_ρ	MRHO	number of ρ values investigated in any computer run
n_e	NE	number of simulated experiments in any computer run
r_1	RETA(1)	number of mean squares pooled for testing at level α_{p2} is integer value of $1 + r_1\hat{\eta}$, fig. 1
r_2	RETA(2)	number of mean squares pooled for testing at level α_f is integer value of $1 + r_2\hat{\eta}$, fig. 1
t	IT	number of treatments
V (. . .)		variance of . . .
$V(e^2)_{max}$	VESQMX	maximum over treatments of sample variance of mean square prediction error over simulations, eq. (35)
x_k		k^{th} independent variable
Y_i		conceptual value of dependent variable
y_i	YOBS(I)	observed value of dependent variable
Z_i	Z(I)	mean squares in Yates' order
α_f		nominal significance level of final test
α_{p1}		nominal significance level of first preliminary test
α_{p2}		nominal significance level of second preliminary test
β_i	B(I)	regression coefficients in Yates' order
δ_k	DELTA(K, M)	parameter determining relative magnitudes of coefficients in population model, eqs. (6) and (8)

Mathematical symbol	FORTRAN name	Description
ζ_j		expectation of j^{th} order statistic of of a $\chi^2_{(1)}$ variable
η		number of mean squares having non-centrality parameter of zero
$\hat{\eta}$	ETA	number of mean squares concluded to be null during any analysis
θ_ℓ	THETA(L)	scale parameter
λ		mean over experiment of noncentrality parameters, eq. (9)
λ_i		noncentrality parameter
μ_i	YMU(I, L, M)	population mean value of Y_i for i^{th} treatment
ρ	KTEM, KRHO(M)	number of non-null coefficients (number of δ_k values) beyond zero order term in population model
$\hat{\rho}$		number of coefficients concluded to be non-null in any analysis iteration
$\bar{\rho}$	AVRHO	mean number of coefficients concluded to be non-null in a strategy investigation
σ		standard deviation of e

APPENDIX B

PROGRAM LISTING

```
C
C     1A.- DECLARATIONS AND TABLES
C
      DIMENSION REMARK(14), ALPHA(11), TB(64,10), RN(16), IND(16),Z(16),
     1THETA(6),  KRHO(3), DELTA(16,3), YMUM(16,3), YMU(16,6,3),
     2AVRHO(6,3), ERSQ(16,6,3), ERSQSQ(16,6,3), ERSQMX(6,3),COERMX(6,3),
     3AVERSQ(6,3), VESQMX(6,3), AVVESQ(6,3), KP(2), RETA(2), ALFA(2)
C
      COMMON KK, YOBS(16), R(16)
C
      DATA (ALPHA(I),I=1,11)/0.001,0.002,0.005,0.01,0.025,0.05,0.10,
     10.25,0.50,0.75,1.0/
      DATA ((TB(I,J),J=1,10),I=1,24)/0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0
     1,0.0,2.0000,1.99999,1.99997,1.99986,1.99917,1.99687,1.9877,1.923,1
     2.706,1.382,2.9976,2.9960,2.9904,2.9809,2.951,2.904,2.806,2.527,2.0
     386,1.688,3.976,3.962,3.925,3.870,3.760,3.625,3.412,2.949,2.395,1.9
     461,4.887,4.845,4.758,4.65,4.44,4.21,3.89,3.287,2.658,2.184,5.74,5.
     563,5.46,5.31,4.99,4.68,4.28,3.57,2.893,2.371,6.51,6.33,6.11,5.67,5
     6.46,5.09,4.61,3.83,3.11,2.54,7.20,6.96,6.65,6.35,5.88,5.44,4.91,4.
     706,3.29,2.69,7.81,7.52,7.10,6.78,6.26,5.75,5.17,4.27,3.45,2.82,8.3
     84,8.01,7.53,7.17,6.59,6.03,5.41,4.45,3.60,2.95,8.82,8.44,7.95,7.53
     9,6.89,6.28,5.61,4.62,3.74,3.07,9.26,8.84,8.33,7.87,7.13,6.50,5.81,
     A4.77,3.87,3.17,9.67,9.21,8.68,8.16,7.37,6.71,5.99,4.92,3.99,3.27,1
     B0.05,9.55,8.95,8.42,7.59,6.91,6.15,5.05,4.10,3.37,10.40,9.86,9.20,
     C8.66,7.79,7.07,6.30,5.17,4.20,3.46,10.72,10.14,9.43,8.83,7.96,7.23
     D,6.44,5.29,4.30,3.55,11.01,10.40,9.64,9.00,8.12,7.38,6.57,5.40,4.3
     E9,3.63,11.28,10.64,9.84,9.17,8.28,7.52,6.69,5.50,4.48,3.70,11.57,1
     F0.86,10.03,9.34,8.43,7.65,6.81,5.60,4.56,3.77,11.76,11.07,10.22,9.
     G51,8.58,7.78,6.92,5.69,4.64,3.84,11.98,11.28,10.40,9.67,8.72,7.90,
     H7.03,5.78,4.71,3.90,12.19,11.48,10.59,9.83,8.86,8.02,7.13,5.87,4.7
     I8,3.96,12.39,11.68,10.76,9.99,8.99,8.10,7.23,5.95,4.85,4.02,12.58,
     J11.87,10.93,10.14,9.12,8.24,7.33,6.03,4.92,4.08/
      DATA ((TB(I,J),J=1,10),I=25,47)/12.76,12.05,11.10,10.29,9.23,8.34,
     17.42,6.11,4.98,4.14,12.93,12.22,11.26,10.43,9.34,8.44,7.51,6.18,5.
     204,4.19,13.09,12.38,11.41,10.56,9.44,8.54,7.60,6.25,5.10,4.24,13.2
     34,12.53,11.55,10.68,9.54,8.63,7.68,6.32,5.16,4.30,13.39,12.68,11.6
     48,10.78,9.64,8.72,7.76,6.38,5.22,4.35,13.53,12.82,11.80,10.88,9.74
     5,8.81,7.83,6.44,5.28,4.40,13.67,12.96,11.91,10.98,9.83,8.89,7.90,6
     6.50,5.33,4.45,13.80,13.09,12.01,11.07,9.91,8.97,7.97,6.56,5.38,4.5
     70,13.93,13.21,12.10,11.16,9.99,9.04,8.04,6.62,5.43,4.54,14.05,13.3
     82,12.19,11.25,10.17,9.11,8.11,6.68,5.48,4.58,14.17,13.43,12.27,11.
     934,10.15,9.18,8.17,6.74,5.53,4.62,14.29,13.53,12.35,11.43,10.22,9.
     A25,8.23,6.80,5.58,4.66,14.41,13.63,12.43,11.51,10.29,9.31,8.29,6.8
```

```
     R5,5.63,4.70,14.53,13.73,12.51,11.59,10.36,9.37,8.35,6.90,5.67,4.74
     C,14.64,13.82,12.59,11.67,10.43,9.43,8.41,6.95,5.71,4.78,14.75,13.9
     D1,12.67,11.75,10.50,9.49,8.46,6.99,5.75,4.82,14.85,14.00,12.75,11.
     E83,10.57,9.55,8.51,7.03,5.79,4.86,14.95,14.09,12.83,11.90,10.64,9.
     F61,8.56,7.07,5.83,4.90,15.05,14.17,12.90,11.97,10.70,9.67,8.61,7.1
     G1,5.87,4.94,15.15,14.25,12.97,12.04,10.76,9.72,8.66,7.15,5.91,4.98
     H,15.24,14.33,13.05,12.11,10.82,9.77,8.71,7.19,5.95,5.01,15.33,14.4
     I0,13.12,12.18,10.88,9.82,8.76,7.23,5.99,5.04,15.42,14.47,13.19,12.
     J25,10.94,9.87,8.81,7.27,6.03,5.07/
      DATA((TB(I,J),J=1,10),I=48,64)/15.50,14.54,13.26,12.32,11.00,9.92,
     18.85,7.31,6.07,5.10,15.58,14.60,13.32,12.38,11.06,9.97,8.89,7.35,6
     2.11,5.13,15.66,14.66,13.38,12.44,11.11,10.02,8.93,7.39,6.14,5.16,1
     35.73,14.72,13.44,12.50,11.16,10.07,8.97,7.43,6.17,5.19,15.80,14.79
     4,13.50,12.56,11.21,10.12,9.01,7.47,6.20,5.22,15.87,14.85,13.56,12.
     562,11.26,10.17,9.05,7.51,6.23,5.25,15.93,14.91,13.62,12.68,11.31,1
     60.21,9.09,7.55,6.26,5.28,15.99,14.97,13.67,12.73,11.36,10.25,9.13,
     77.59,6.29,5.31,16.05,15.03,13.72,12.78,11.40,10.29,9.17,7.63,6.32,
     85.34,16.11,15.10,13.77,12.83,11.44,10.33,9.21,7.67,6.35,5.37,16.17
     9,15.16,13.82,12.88,11.48,10.37,9.25,7.70,6.38,5.40,16.23,15.22,13.
     A87,12.93,11.52,10.41,9.29,7.73,6.41,5.43,16.29,15.28,13.92,12.97,1
     B1.56,10.45,9.33,7.76,6.44,5.46,16.34,15.34,13.97,13.01,11.60,10.49
     C,9.37,7.79,6.47,5.48,16.39,15.40,14.02,13.05,11.64,10.53,9.41,7.82
     D,6.50,5.50,16.44,15.46,14.06,13.09,11.67,10.57,9.45,7.85,6.53,5.52
     E,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0/
C
      NAMELIST /OUT1/ M /OUT2/   YOBS    /OUT3/   YMUM       /OUT4/ M,L
      NAMELIST /OUT5/   YMU        /OUT6/ N /OUT7/   RN
      NAMELIST /OUT8/ J,INDX
      NAMELIST /OUT9/ MPX,KPX,KF
      NAMELIST /OUT 10/ B
      NAMELIST /OUT11/ IND,Z,B
      NAMELIST /OUT13/ N,M,L           /OUT14/J,JN,TEM /OUT15/ ETA
      NAMELIST /OUT16/ NC
      NAMELIST /                    /OUT20/ ETA,AVRHO
      NAMELIST /                 /OUT22/   ERSQ       ,ERSQSO
C
C     10.- INPUTS AND CONSTANTS
C
      READ(5,800) (REMARK(I),I=1,14)
      WRITE(6,801) (REMARK(I),I=1,14)
      READ(5,802) LGMH, NE, MRHO, KODE
      KG= LGMH
      IT= 2**LGMH
      WRITE(6,803) KG, IT, MRHO, NE
      ITM1= IT-1
      ITM2 = IT-2
      FIT= IT
      FITM1 = ITM1
      FNE= NE
      IF (NE.EQ.0) FNE = 1.0
      NEM1 = NE - 1
      FNEM1 = NEM1
      IF (NEM1 .EQ. 0) FNEM1 = 1.0
      READ(5,804) LTH, (THETA(L), L=1,LTH)
      DO 5  M=1,MRHO
      READ(5,806) KTEM,(DELTA(K,M), K=1,KTEM)
      KRHO(M)= KTEM
```

```
      WRITE(6,807)  KTEM, (DELTA(K,M), K=1,KTEM)
    5 CONTINUE
      IP= 2**KG
      FIP = IP
      IPP1= IP + 1
      FENBIP = FNE*FIP
      FEM1IP = FNEM1*FIP
C
C     1C.- POPULATION MEANS
C
      KK= KC
      DO 40  M=1,MRHO
      INDXM = M
      DO 10 I=1,IP
      B(I) = 0.0
   10 CONTINUE
      KTEM= KRHO(M)
      KTEMP2 = KTEM + 2
      DO 20  I=1,KTEM
      B(I+1)= DELTA(I,M)
   20 CONTINUE
      IF (KODE .GT. 2 ) WRITE (6, OUT1 )
      DO 22  I=1,IP
      IPP1MI= IPP1 - I
      YOBS(I)= B(IPP1MI)
   22 CONTINUE
      IF (KODE .GT. 1 ) WRITE (6, OUT2 )
      CALL YATES
      DO 30  I=1,IP
      IPP1MI= IPP1 - I
      YMUM(I,M)= B(IPP1MI)
   30 CONTINUE
   40 CONTINUE
      IF (KODE .GT. 2 ) WRITE (6, OUT3 )
      DO 49  M=1,MRHO
      INDXM = M
      DO 48  L=1,LTH
      INDXL = L
      IF (KODE .GT. 3 ) WRITE (6, OUT4 )
      DO 47  I=1,IP
      YMU(I,L,M) = YMUM(I,M)*THETA(L)
   47 CONTINUE
   48 CONTINUE
   49 CONTINUE
      IF (KODE .GT. 4 ) WRITE (6, OUT5 )
C
C     1D. - STRATEGY
C
   50 READ (5,808,END = 899) MP,KP(1),KP(2), KF,RETA(1),RETA(2)
      DO 55 NC=1,2
      KPX= KP(NC)
      ALFA(NC)= ALPHA(KPX)
   55 CONTINUE
      CALL SAND (XS)
      DO 99  M=1,MRHO
      DO 98 L=1,LTH
      AVRHO(L,M)= 0.0
```

```
      DO 97  I=1,IP
      ERSQ(I,L,M)= 0.0
      EPSCSQ(I,L,M) = 0.0
   97 CONTINUE
   98 CONTINUE
   99 CONTINUE
C
C     2.- SIMULATIONS AND MODEL FITTING
C
      DO 699  N=1,NE
      INDXN = N
      IF (KODE .GT. 5 ) WRITE (6, OUT6 )
      IF (NF.EQ.0) GO TO 200
      DO 213 I=1,IT
      CALL RAND(RN(I))
  213 CONTINUE
      IF (KODE .GT. 6 ) WRITE (6, OUT7 )
      DO 215 I=1,IT,2
      E= SQRT(-2.0*ALOG(RN(I)))
      D= 6.2831853*RN(I+1)
      RN(I)= E*COS(D)
      RN(I+1)= E*SIN(D)
  215 CONTINUE
      IF (KODE .GT. 6 ) WRITE (6, OUT7 )
      GO TO 201
  200 DO 209 I=1,IT
      RN(I) = 0.0
  209 CONTINUE
  201 DO 695  M=1,MRHO
      INDXM = M
      DO 680 L=1,LTH
      INDXL = L
      KK= LGMH
      IF (KODE .GT. 12) WRITE (6, OUT13)
  204 DO 214  I=1,IT
      YOBS(I)= YMU(I,L,M) +RN(I)
  214 CONTINUE
      IF (KODE .GT. 1 ) WRITE (6, OUT2 )
      CALL YATES
      GO TO 300
C
C     3.- CONSTRUCTION AND ORDERING OF MEAN SQUARES
C
  300 DO 309  I=1,IT
      IND(I)= I
      Z(I)= B(I+1)*B(I+1)/FIT
      B(I) = B(I) / FIT
  309 CONTINUE
      IF (KODE .GT. 10) WRITE (6, OUT11)
      IF (MP .LT. 1) GO TO 432
C
      DO 313  J=1,ITM2
      TEST= Z(ITM1)
      IN= ITM1
      DO 312  NA=J,ITM2
      IF(TEST-Z(NA))  312,312,311
  311 TEST = Z(NA)
```

```
      IN= NA
  312 CONTINUE
      ITEM= IND(IN)
      TEM= Z(IN)
      IND(IN)= IND(J)
      Z(IN)= Z(J)
      IND(J)= ITEM
      Z(J) = TEM
  313 CONTINUE
      IF (KODE .GT. 10) WRITE (6, OUT11)
C
C     4.- DELETION OF TERMS
C
      MPX= MP
      IF (KODE .GT. 12) WRITE (6, OUT13)
      DO 424  NC = 1,2
      INDXNC = NC
      IF (KODE .GT. 15) WRITE (6, OUT16)
      KPX = KP(NC)
      IF (KODE .GT. 8 ) WRITE (6, OUT9 )
      IF (KPX .EQ. 11) GO TO 424
      MPP1= MPX + 1
      JN= MPP1
      TEM= 0.0
      DO 415  J = 1,MPX
      INDXJ = J
      TEM = TEM + Z(J)
      IF (KODE .GT. 13) WRITE (6, OUT14)
  415 CONTINUE
      DO 419 J=MPP1,ITM1
      INDXJ = J
      FJN = JN
      TEST = FJN*Z(J)/(TEM + Z(J))
      IF (TEST - TR(JN,KPX)) 416,416,420
  416 TEM = TEM + Z(J)
      JN = JN + 1
      IF (KODE .GT. 13) WRITE (6, OUT14)
  419 CONTINUE
      JETA = ITM1
      GO TO 422
  420 JETA = J-1
  422 ETA = JETA
      IF (KODE .GT. 14) WRITE (6, OUT15)
      IF (KODE .GT. 12) WRITE (6, OUT13)
      IF (RETA(NC) .LT. 1.0) GO TO 421
      MPX = JETA
      GO TO 424
  421 MPX = 1 + IFIX(RETA(NC)*ETA)
  424 CONTINUE
      IF (KODE .GT. 8 ) WRITE (6, OUT9 )
      MPP1   = MPX + 1
      JN = MPP1
      TEM = 0.0
      DO 425  J=1,MPX
      INDXJ = J
      INDX = IND(J)+1
      B(INDX)= 0.0
```

```
      TEM = TEM + Z(J)
      IF (KODE .GT. 7 ) WRITE (6, OUT8 )
      IF (KODE .GT. 9 ) WRITE (6, OUT10)
  425 CONTINUE
      IF ( KF .EQ. 11 ) GO TO 434
      DO 429  J=MPP1,ITM1
      INDXJ = J
      FJN = JN
      TEST = FJN*Z(J)/(TEM + Z(J))
      IF (TEST - TB(JN,KF)) 428,428,430
  428 INDX = IND(J) + 1
      IF (KODE .GT. 7 ) WRITE (6, OUT8 )
      B(INDX) = 0.0
      IF (KODE .GT. 9 ) WRITE (6, OUT10)
  429 CONTINUE
      ETA = FITM1
      GO TO 433
  430 JETA = J-1
      ETA = JETA
      IF ( KODE .LT. 8 ) GO TO 433
      IF (KODE .GT. 14) WRITE (6, OUT15)
      DO 431  J=JETA,IT
      INDXJ = J
      INDX = IND(J)+1
      IF (KODE .GT. 7 ) WRITE (6, OUT8 )
  431 CONTINUE
      IF (KODE .GT. 9 ) WRITE (6, OUT10)
      GO TO 433
  432 ETA = 0.0
      GO TO 433
  434 ETA = MPX
  433 AVRHO(L,M) = AVRHO(L,M) + FITM1 - ETA
C
C     5.- PREDICTIONS
C
  500 KK = KG
  540 DO 546 I=1,IP
      IPP1MI = IPP1-I
      YOBS(I) = B(IPP1MI)
  546 CONTINUE
      IF (KODE .GT. 1 ) WRITE (6, OUT2 )
      CALL YATES
      IF (KODE .GT. 9 ) WRITE (6, OUT10)
      GO TO 600
C
C     6.- ACCUMULATION OF ERRORS
C
  600 DO 609  I=1,IP
      IPP1MI = IPP1-I
      TEM= (B(IPP1MI) - YMU(I,L,M))**2
      ERSQ(I,L,M) = ERSQ(I,L,M) + TEM
      ERSQSQ(I,L,M)= ERSQSQ(I,L,M) + TEM**2
  609 CONTINUE
  680 CONTINUE
  690 CONTINUE
      IF (KODE .GT. 19) WRITE (6, OUT20)
      IF (KODE .GT. 21) WRITE (6, OUT22)
```

```
      IF (NE .EQ. 0) GO TO 700
  699 CONTINUE
C
C     7.- DETERMINATION OF MAXIMUM AND MEAN SQUARED ERRORS
C
  700 DO 790 M=1,MRHO
      DO 780 L=1,LTH
      C = 0.0
      D = 0.0
      E= 0.0
      F= 0.0
      DO 750 I=1,IP
      C = AMAX1(C,ERSQ(I,L,M))
      D = D + ERSQ(I,L,M)
      TEM= ERSQSQ(I,L,M)- ((ERSQ(I,L,M))**2)/FNE
      E = AMAX1(E,TEM)
      F = F + TEM
  750 CONTINUE
      ERSQMX(L,M) = C/FNE
      COERMX(L,M) = (SQRT(ERSQMX(L,M)))/THETA(L)
      AVERSQ(L,M) = D/FENBIP
      IF (NE .EQ. 0) GO TO 780
      VESQMX(L,M)= E/FNEM1
      AVVESQ(L,M)= F/FEM1TP
      AVRHO(L,M) = AVRHO(L,M)/FNE
  780 CONTINUE
  790 CONTINUE
C
C     8.- OUTPUT
C
      WRITE (6,809) (MP, ALFA(1), RETA(1), ALFA(2), RETA(2), ALPHA(KF))
      WRITE (6,811) (KRHO(M), M=1,MRHO)
      WRITE (6,813)
      WRITE (6,815)
      WRITE (6,817) (THETA(L),(AVRHO (L,M),M=1,MRHO),L=1,LTH)
      WRITE (6,831)
      WRITE (6,817) (THETA(L),(ERSQMX(L,M),M=1,MRHO),L=1,LTH)
      WRITE (6,835)
      WRITE (6,817) (THETA(L),(AVERSQ(L,M),M=1,MRHO),L=1,LTH)
      WRITE(6,836)
      WRITE (6,817) (THETA(L),(VESQMX(L,M),M=1,MRHO),L=1,LTH)
      WRITE(6,837)
      WRITE (6,817) (THETA(L),(AVVESQ(L,M),M=1,MRHO),L=1,LTH)
      WRITE (6,838)
      WRITE (6,817) (THETA(L),(COFRMX(L,M),M=1,MRHO),L=1,LTH)
      GO TO 50
  899 STOP
C
  800 FORMAT (13A6,A2)
  801 FORMAT (1H1,//10X,13A6,A2//)
  802 FORMAT (4I5)
  803 FORMAT   (1H0,3X,4HKG =I5,5X,4HIT =I5,5X,6HMRHO =I5,5X,4HNE =I5)
  804 FORMAT (I8,6F8.3)
  806 FORMAT (I4/(10F8.5))
  807 FORMAT (1H0,5HRHO =I5,5X,7HDELTA =//(1X,10F10.5))
  808 FORMAT (4I5, 2F5.3)
  809 FORMAT (1H1//,1X,3HMP=I5,5X,9HALPHAP1 =F6.3,5X,7HRETA1 =F6.3,5X,
```

```
     19HALPHAP2 =F6.3,5X,7HRETA2 =F6.3,5X,8HALPHAF =F6.3)
 811 FORMAT (1H0,5HRHO =3I13//)
 813 FORMAT (1H0,5HTHETA)
 815 FORMAT  (1H0,20X,5HAVRHO//)
 817 FORMAT (1X,F8.3,3E14.4)
 831 FORMAT (1H0,20X,6HERSQMX//)
 835 FORMAT (1H0,20X,6HAVERSQ//)
 836 FORMAT (1H0,20X,6HVESQMX//)
 837 FORMAT (1H0,20X,6HAVVESQ//)
 838 FORMAT (1H0,20X,6HCOERMX//)
     END

      SUBROUTINE YATES
C
C     9.- YATES METHOD SUBROUTINE
C
      COMMON KK,Y(16),B(16)
      II = 2**KK
      IIDB2 = II/2
      KKM1 = KK-1
      DO 908 K=1,KKM1
      DO 906 I=1,II,2
      IP1D2 = (I+1)/2
      B(IP1D2) = Y(I+1)+Y(I)
      LL = IP1D2+IIDB2
  906 B(LL) = Y(I+1)-Y(I)
      DO 907 I=1,II
  907 Y(I) = B(I)
  908 CONTINUE
      DO 909 I=1,II,2
      IP1D2 = (I+1)/2
      B(IP1D2) = Y(I+1)+Y(I)
      LL = IP1D2+IIDB2
      B(LL) = Y(I+1)-Y(I)
  909 CONTINUE
      RETURN
      END
```

APPENDIX C

ILLUSTRATIVE OUTPUT

```
AUG 24, 1977  IF (NEM1 WAS IF (FNEM1   IF (RETA SUBSCRIPTED JULY 29, 1977

   K6 =     4     IT =   16     MRHO =     3     NE = 1000

RHO =   11     DELTA =

   1.98710   1.51950   1.24640   1.04300    .87550    .72980    .59850    .47710    .36260    .25260
    .14410

RHO =   13     DELTA =

   2.05280   1.59840   1.33540   1.14090    .98200    .84470    .72200    .60950    .50440    .40480
    .30930    .21630    .12370

RHO =   15     DELTA =

   2.10819   1.66452   1.40939   1.22186   1.06945    .93855    .82213    .71606    .61764    .52503
    .43685    .35211    .26985    .18921    .10835
```

MP= 1 ALPHAP1 = 1.000 BETA1 = .000 ALPHAP2 = .750 BETA2 = .360 ALPHAF = 1.000

RHO = 11 13 15

THETA

AVRHO

THETA	11	13	15
.125	.1366+02	.1371+02	.1371+02
.250	.1368+02	.1370+02	.1372+02
.500	.1370+02	.1375+02	.1367+02
1.000	.1379+02	.1375+02	.1368+02
2.000	.1381+02	.1385+02	.1372+02

ERSQMX

THETA	11	13	15
.125	.1112+01	.1112+01	.1111+01
.250	.1105+01	.1106+01	.1110+01
.500	.1100+01	.1116+01	.1112+01
1.000	.1099+01	.1096+01	.1120+01
2.000	.1116+01	.1112+01	.1273+01

AVERSQ

THETA	11	13	15
.125	.1005+01	.1005+01	.1005+01
.250	.1006+01	.1007+01	.1007+01
.500	.1008+01	.1010+01	.1018+01
1.000	.1006+01	.1010+01	.1045+01
2.000	.1004+01	.1011+01	.1129+01

VESQMX

THETA	11	13	15
.125	.2309+01	.2284+01	.2276+01
.250	.2290+01	.2318+01	.2301+01
.500	.2302+01	.2317+01	.2401+01
1.000	.2289+01	.2310+01	.2513+01
2.000	.2306+01	.2335+01	.382[illegible]+01

AVVESO

.125	.1994+01	.1986+01	.1992+01
.250	.1988+01	.1992+01	.1997+01
.500	.2006+01	.2015+01	.2055+01
1.000	.2000+01	.2004+01	.2165+01
2.000	.1980+01	.2024+01	.2587+01

COERMX

.125	.8434+01	.8438+01	.8432+01
.250	.4205+01	.4237+01	.4214+01
.500	.2098+01	.2113+01	.2109+01
1.000	.1048+01	.1047+01	.1058+01
2.000	.5282+00	.5272+00	.5642+00

BIBLIOGRAPHY

Allen, D. M. (1971) Mean Square Error of Prediction as a Criterion for Selecting Variables. Technometrics 13. 469-475, discussion 477-481.

Amling, G. E., and Holms, A. G. (1973) POOLMS - A Computer Program for Fitting and Model Selection for Two-Level Factorial Replication-Free Experiments. NASA TM X-2706.

Birnbaum, A. (1959)· On the Analysis of Factorial Experiments Without Replication. Technometrics 1 343-357.

Daniel, C. (1959). Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments. Technometrics 1 311-341.

Davies, O. L., ed. (1956). The Design and Analysis of Industrial Experiments. 2nd. ed. rev. Hafner, New York.

Duckworth, W. E. (1965)· Statistical Method in Metallurgical Development. The Statistician 15: (1) 7-30.

Harter, H. L. (1964): Expected Values of Exponential, Weibull, and Gamma Order Statistics. ARL 64-31, Aerospace Research Laboratories, Applied Mathematics Research Laboratory, Wright-Patterson Air Force Base, Ohio. (Available from DDC as AD-436 763.)

Holms, A. G. (1966): Multiple-Decision Procedures for the ANOVA of Two-Level Factorial Replication-Free Experiments. PhD Thesis, Western Reserve University, June 1966. (Available from University Microfilms, Inc., Ann Arbor, Mich.)

Holms, A. G., and Berrettoni, J. N. (1967)· Multiple Decision Procedures for ANOVA of Two-Level Factorial Fixed Effects Replication-Free Experiments. NASA TN D-4272.

Holms, A. G., and Berrettoni, J. N. (1969). Chain-Pooling ANOVA for Two-Level Factorial Replication-Free Experiments. Technometrics 11. 725-46.

Holms, Arthur G. (1974). Chain Pooling to Minimize Prediction Error in Subset Regression. NASA TM X-71645.

Holms, Arthur G. (1977). Concepts for the Development of Light-Weight Composite Structures for Rotor Burst Containment. NASA TM X-73633.

Kendall, M. G. and Stuart, A. (1961): The Advanced Theory of Statistics, Vol. 2. Hafner, New York.

Kennedy, W. J., and Bancroft, T. A. (1971). Model Building for Prediction in Regression Based Upon Repeated Significance Tests. Ann. Math. Stat. 42 1273-1284.

Mangano, G. J., Salvino, J. T., and DeLucia, R. A. (1977). Rotor Burst Protection Program. Experimentation to Provide Guidelines for the Design of Turbine Rotor Burst Fragment Containment Rings. NASA CR-135166, Naval Air Propulsion Test Center, Trenton, N.J. NAPTC-PE-98.

Sidik, Steven M. (1972): An Improved Multiple Linear Regression and Data Analysis Computer Program Package. NASA TN D-6770.

Walls, R. C., and Weeks, D. L. (1969). A Note on the Variance of a Predicted Response in Regression. Am. Stat. 23: (3) 23-26.

Wilk, M. B., Gnanadesikan, R., and Freeny, A. E. (1963) Estimation of Error Variance from Smallest Ordered Contrasts. J. Am. Stat. Assoc. 58: 152-160.

TABLE I

INVESTIGATED VALUES OF SCALE PARAMETER AND ASSOCIATED VALUES OF MEAN NONCENTRALITY PARAMETER

Scale parameter, θ	Mean noncentrality parameter, λ
0.125	0.25
.250	1.00
.500	4.00
1.000	16.00
2.000	64.00

TABLE II

UNFAVORABLE DISTRIBUTIONS OF RELATIVE VALUES OF MODEL COEFFICIENTS

J	ρ					
	15		13		11	
	$\frac{1}{2}\zeta_J$	δ_1	$\frac{1}{2}\zeta_J$	δ_1	$\frac{1}{2}\zeta_J$	δ_1
1	0.00587	0.10835	0.00765	0.1237	0.01038	0.1441
2	.01790	.18921	.02339	.2163	.03190	.2526
3	.03641	.26985	.04782	.3093	.06574	.3626
4	.06199	.35211	.08195	.4048	.11382	.4771
5	.09542	.43685	.12721	.5044	.17912	.5985
6	.13783	.52503	.18572	.6095	.26634	.7298
7	.19074	.61764	.26062	.7220	.38328	.8755
8	.25637	.71606	.35677	.8447	.54390	1.0430
9	.33795	.82213	.48212	.9820	.77674	1.2464
10	.44044	.93855	.65081	1.1409	1.15440	1.5195
11	.57186	1.06945	.89158	1.3354	1.97435	1.9871
12	.74647	1.22186	1.27744	1.5984	-------	------
13	.99319	1.40939	2.10691	2.0528	-------	------
14	1.38532	1.66452	-------	------	-------	------
15	2.22223	2.10819	-------	------	-------	------

TABLE III

PARAMETER COMBINATIONS, $\beta_k = \theta\,\delta_k$, $\rho = 15$

k	δ_k	θ				
		0.125	0.250	0.500	1.000	2.000
1	2.1082	0.2635	0.5270	1.0541	2.1082	4.2164
2	1.6645	.2081	.4161	.8322	1.6645	3.3290
3	1.4094	.1762	.3524	.7047	1.4094	2.8188
4	1.2219	.1527	.3055	.6110	1.2219	2.4438
5	1.0694	.1337	.2674	.5347	1.0694	2.1388
6	.9386	.1173	.2346	.4693	.9386	1.8772
7	.8221	.1028	.2205	.4110	.8221	1.6442
8	.7161	.0895	.1790	.3580	.7161	1.4322
9	.6176	.0772	.1544	.3088	.6176	1.2352
10	.5250	.0656	.1312	.2625	.5250	1.0500
11	.4368	.0546	.1092	.2184	.4368	.8736
12	.3521	.0440	.0880	.1760	.3521	.7042
13	.2699	.0337	.0675	.1350	.2699	.5398
14	.1892	.0237	.0473	.0946	.1892	.3784
15	.1084	.0136	.0271	.0542	.1084	.2168

$\sum_{k=1}^{15} \delta_k$: 12.4492

$12.4492\,\theta$	1.556	3.112	6.225	12.449	24.898
$(12.4492\,\theta)^{-1}$.	0.643	0.321	0.161	0.080	0.040

TABLE IV

VALUES OF $C_{ee,mx}(\theta)$ FOR STRATEGIES NEIGHBORING BETTER STRATEGIES AT $\theta = 0\ 125$

[$\alpha_{p1} = 1\ 00$, $\alpha_f = 1\ 00$, $\rho = 15$, $n_e = 1000$]

m_p	α_{p2}	r_2									
		0 650		0 675		0 700		0 725		0 750	
		θ									
		0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000
1	0 001	7 759	4 964	7 730	6 038	7 730	6 038	7 731	6 038	7 961	7 263
	002	7 760	4 957	7 730	6 029	7 730	6 029	7 731	6 020	7 960	7 252
	005	7 751	4 942	7 718	6 011	7 719	6 011	7 723	6 011	7 943	7 230
	01	7 758	4 933	7 740	5 999	7 741	5 999	7 745	5 999	7 986	7 216
	025	7 765	4 882	7 762	5 938	7 764	5 938	7 761	5 938	7 992	7 142
	05	7 792	4 802	7 851	5 840	7 854	5 840	7 870	5 840	8 049	7 024
2	0 001	7 759	4 964	7 730	6 038	7 730	6 038	7 731	6 038	7 961	7 263
	002	7 760	4 962	7 730	6 035	7 730	6 035	7 730	6 035	7 962	7 259
	005	7 753	4 956	7 725	6 029	7 726	6 029	7 729	6 029	7 949	7 251
	01	7 757	4 954	7 742	6 026	7 744	6 026	7 748	6 026	7 972	7 248
	025	7 761	4 932	7 732	6 000	7 733	6 000	7 730	6 000	7 952	7 217
	05	7 775	4 901	7 820	5 961	7 824	5 961	7 840	5 961	8 013	7 170
3	0 002	7 762	4 964	7 719	6 038	7 719	6 038	7 719	6 038	7 952	7 263
	005	7 763	4 962	7 714	6 035	7 715	6 035	7 718	6 035	7 941	7 259
	01	7 761	4 959	7 717	6 032	7 718	6 032	7 722	6 032	7 955	7 256
	025	7 774	4 954	7 710	6 026	7 711	6 026	7 708	6 026	7 941	7 248
	05	7 787	4 942	7 739	6 011	7 743	6 011	7 760	6 011	7 956	7 230
4	0 005	7 768	4 964	7 712	6 038	7 713	6 038	7 716	6 038	7 939	7 263
	01	7 767	4 964	7 711	6 038	7 713	6 038	7 717	6 038	7 950	7 263
	025	7 778	4 964	7 699	6 038	7 701	6 038	7 697	6 038	7 933	7 263
	05	7 786	4 962	7 706	6 035	7 709	6 035	7 726	6 035	7 939	7 259
	10	7 812	5 950	7 818	6 021	7 813	6 021	7 804	6 021	8 005	7 242
5	0 005	7 768	4 964	7 712	6 038	7 713	6 038	7 716	6 038	7 939	7 263
	01	7 767	4 964	7 717	6 038	7 719	6 038	7 723	6 038	7 953	7 263
	025	7 772	4 964	7 711	6 038	7 712	6 038	7 709	6 038	7 937	7 263
	05	7 774	4 964	7 689	6 038	7 691	6 038	7 708	6 038	7 920	7 263
	10	7 800	4 962	7 786	6 036	7 777	6 036	7 770	6 036	7 979	7 259
6	0 01	7 767	4 964	7 717	6 038	7 719	6 038	7 723	6 038	7 953	7 263
	025	7 772	4 964	7 712	6 038	7 714	6 038	7 710	6 038	7 936	7 263
	05	7 772	4 964	7 699	6 038	7 701	6 038	7 719	6 038	7 939	7 263
	10	7 791	4 964	7 788	6 038	7 779	6 038	7 774	6 038	7 995	7 263

TABLE V

ADDITIONAL PROPERTIES FOR STRATEGIES THAT WERE GOOD AT $\theta = 0.125$

[$\alpha_{p1} = 1.00$, $\alpha_f = 1.00$, $\rho = 15$, $n_e = 1000$.]

m_p	α_{p2}	r_2	θ					
			0.125	2.000	0.125	2.000	0.125	2.000
			$\bar{\rho}$		$V(e^2)_{mx}$		$C_{ee,mx}(\theta)$	
1	0.005	0.675	4.079	4.086	2.025	372.4	7.718	6.011
2	.005	.675	4.058	4.026	2.026	248.2	7.725	6.029
3	.025	.725	4.152	4.032	1.996	268.4	7.708	6.026
4	.025	.725	4.112	4.000	1.995	185.4	7.697	6.038
5	.05	.675	4.225	4.000	1.973	185.4	7.689	6.038
6	.05	.675	4.171	4.000	1.980	185.4	7.699	6.038

TABLE VI

VALUES OF $C_{ee,mx}(\theta)$ FOR STRATEGIES WITH $m_p = 1$, $\alpha_{p1} = 1\ 00$, AND $\alpha_f = 1\ 00$

r_2	α_{p2}																			
	0 75		0 50		0 25		0 10		0 05		0 025		0 01		0 005		0 002		0 001	
	θ																			
	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000
0 100	8 444	0 5339	8 459	0 5431	8 447	0 5687	8 449	0 5800	8 445	0 5829	8 446	0 5829	8 446	0 5842	8 444	0 5845	8 444	0 5849	8 444	0 5850
200	8 456	5339	8 458	5713	8 430	6483	8 433	6815	8 427	6939	8 423	6990	8 424	7012	8 423	7014	8 422	7022	8 422	7031
300	8 448	6410	8 439	8013	8 383	1 153	8 322	1 295	8 286	1 338	8 265	1 355	8 262	1 370	8 257	1 372	8 258	1 375	8 258	1 375
400	8 434	5786	8 386	1 033	8 312	1 539	8 257	1 758	8 215	1 823	8 195	1 850	8 188	1 870	8 183	1 873	8 186	1 877	8 185	1 879
500	8 423	6613	8 361	1 573	8 275	2 566	8 167	2 977	8 139	3 096	8 111	3 146	8 076	3 180	8 077	3 186	8 077	3 196	8 078	3 199
600	8 417	7277	8 349	2 083	8 070	3 926	7 943	4 613	7 801	4 802	7 769	4 882	7 758	4 933	7 750	4 942	7 757	4 957	7 756	4 964
700	8 411	8376	8 304	2 586	7 985	4 770	7 930	5 608	7 854	5 840	7 764	5 938	7 741	5 999	7 719	6 011	7 730	6 029	7 730	6 038
800	8 399	9670	8 206	3 386	8 073	5 772	8 086	6 745	8 052	7 024	7 990	7 142	7 985	7 216	7 948	7 230	7 956	7 252	7 957	7 263

TABLE VII

PROPERTIES OF STRATEGIES ADMISSIBLE FOR $\theta = 0\ 125$ AND FOR $\theta = 2\ 000$

[$R(0\ 125) = C_{ee,mx}/7\ 689$, $R(2\ 000) = C_{ee,mx}/0\ 5282$, $\rho = 15$, $n_e = 1000$]

m_p	α_{p1}	r_1	α_{p2}	r_2	α_f	θ							
						0 125	2 000	0 125	2 000	0 125	2.000	0 125	2 000
						$\bar{\rho}$		$V(e^2)_{mx}$		$C_{ee,mx}$		$R(\theta)$	
0	----	-----	----	-----	----	15 00	15 00	2 282	2 282	8 451	0 5282	1 0991	1 0000
1	1 00	-----	0 75	0 075	1 00	13 99	14 00	2 299	2 615	8 448	5316	1 0987	1 0064
1	1 00	-----	75	125	1 00	13 96	13 97	2 294	2 669	8 446	5327	1 0985	1 0085
1	1 00	-----	75	077	1 00	13 98	13 99	2 298	2 617	8 445	5330	1 0983	1 0091
1	1 00	-----	75	084	1 00	13 98	13 99	2 298	2 633	8 444	5337	1 0982	1 0104
3	1 00	-----	75	200	1 00	13 70	13 71	2 296	2 968	8 441	5477	1 0978	1 0369
1	1 00	-----	75	350	1 00	13 72	13 72	2 275	3 477	8 435	5612	1 0970	1 0625
1	75	0 350	75	350	1 00	13 72	13 72	2 275	3 477	8 435	5612	1 0970	1 0625
1	1 00	-----	75	360	1 00	13 71	13 72	2 276	3 820	8 432	5642	1 0966	1 0682
1	1 00	-----	75	410	1 00	13 67	13 68	2 288	5 226	8 426	5829	1 0959	1 1036
2	1 00	-----	75	350	1 00	13 32	13 38	2 281	5 038	8 424	6136	1 0956	1 1617
2	1 00	-----	75	400	1 00	13 27	13 34	2 280	7 758	8 422	6442	1 0953	1 2196
4	1 00	-----	75	300	1 00	12 51	12 58	2 240	5 800	8 414	6660	1 0943	1 2609
2	1 00	-----	75	450	1 00	13 16	13 23	2 300	12 52	8 413	7003	1 0942	1 3258
3	1 00	-----	75	400	1 00	12 49	12 57	2 257	11 32	8 406	7269	1 0932	1 3762
1	1 00	-----	75	400	75	13 25	13 26	2 301	60 36	8 405	7737	1 0931	1 4648
1	75	700	75	400	75	13 25	13 26	2 301	60 36	8 405	7737	1 0931	1 4648
2	1 00	-----	75	300	75	12 81	12 91	2 294	24 20	8 400	7945	1 0925	1 5042
2	1 00	-----	75	500	1 00	12 39	12 48	2 299	27 70	8 397	7995	1 0921	1 5136
4	1 00	-----	75	400	1 00	12 24	12 27	2 243	16 17	8 375	8204	1 0892	1 5532
3	1 00	-----	75	300	75	12 41	12 52	2 283	31 87	8 356	8912	1 0867	1 6872

TABLE VII - Continued

m_p	α_{p1}	r_1	α_{p2}	r_2	α_f	θ							
						0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000
						$\overline{\rho}$		$V(e^2)_{mx}$		$C_{ee,mx}$		$R(\theta)$	
1	1 00	-----	0 025	0 250	1 00	11 12	11 10	2 305	11 40	8 340	0 9653	1 0847	1 8275
1	1 00	-----	005	250	1 00	11 02	11 03	2 306	11 45	8 339	9760	1 0845	1 8478
4	1 00	-----	75	300	75	11 64	11 82	2 243	39 26	8 328	9846	1 0831	1 8641
4	1 00	-----	75	500	1 00	11 13	11 20	2 209	68 70	8 290	1 169	1 0782	2 213
4	1 00	-----	75	400	75	11 28	11 36	2 207	159 9	8 288	1 310	1 0779	2 480
1	1 00	-----	05	325	1 00	10 40	10 26	2 265	22 70	8 285	1 338	1 0775	2 533
1	1 00	-----	025	325	1 00	10 16	10.13	2 260	21 81	8 264	1 355	1 0748	2 565
1	1 00	-----	01	300	1 00	10 07	10 05	2 256	21 35	8 262	1 370	1 0745	2 594
1	05	0 500	01	300	1 00	10 07	10.05	2 256	21 35	8 262	1 370	1 0745	2 594
1	1 00	-----	005	300	1 00	10 03	10 04	2 261	21 20	8 257	1 372	1 0739	2 598
5	1 00	-----	50	400	1 00	10 72	10 11	2 257	55 52	8 241	1 594	1 0718	3 018
5	1 00	-----	75	600	1 00	9 883	10 09	2 270	211 0	8 226	1 647	1 0698	3 118
4	1 00	-----	25	400	1 00	10 01	9 286	2 237	43 37	8 213	1 820	1 0681	3 446
1	1 00	-----	025	400	1 00	9 202	9 161	2 215	39 96	8 195	1 850	1 0658	3 502
1	025	400	025	400	1 00	9 202	9 161	2 215	39 96	8 195	1 850	1 0658	3 502
1	1 00	-----	01	400	1 00	9 077	9 063	2 213	37 37	8 188	1 870	1 0649	3 540
1	1 00	-----	005	400	1 00	9 034	9 044	2 218	36 77	8 183	1 873	1 0642	3 546
4	1 00	-----	75	600	75	9 794	9 889	2 218	752 0	8 170	2 036	1 0626	3 855
4	1 00	-----	75	800	1 00	9 612	9 748	2 296	1298 0	8 150	2 138	1 0600	4 048
5	1 00	-----	50	400	75	9 577	9 106	2 100	346 1	8 143	2 250	1 0591	4 260
5	1 00	-----	75	600	75	8 859	9 134	2 167	910 9	8 140	2 337	1 0587	4 424
1	05	500	05	300	50	8 589	8 553	2 118	385 0	8 080	2 436	1 0509	4 612
2	05	700	05	300	50	8 398	8 333	2 140	369 8	8 059	2 474	1 0481	4 684
1	005	800	01	300	50	8 159	8 267	2 140	365 6	8 031	2 491	1 0445	4 716
1	1 00	-----	025	400	50	7 250	7 281	1 924	701 6	7 989	3 234	1 0390	6 123
1	025	800	025	400	50	7 250	7 281	1 924	701 6	7 989	3 234	1 0390	6 123

TABLE VII - Concluded

m_p	α_{p1}	r_1	α_{p2}	r_2	α_f	θ							
						0 125	2 000	0 125	2 000	0 125	2 000	0 125	2 000
						$\bar{\rho}$		$V(e^2)_{mx}$		$C_{ee,mx}$		$R(\theta)$	
2	1 00	-----	0 025	0 400	0 50	7 152	7 143	1 906	677 4	7 971	3 262	1 0367	6 176
2	025	0 800	025	400	50	7 152	7 143	1 906	677 4	7 971	3 262	1 0367	6 176
1	50	600	005	400	50	7 057	7 131	1 907	670 4	7 960	3 270	1 0352	6 191
2	01	825	01	350	50	7 092	7 097	1 911	665 8	7 958	3 276	1 0350	6 202
2	01	825	.025	375	50	7 092	7 097	1 908	665 8	7 958	3 276	1 0350	6 202
1	05	800	01	300	25	6 901	6 798	1 920	1323 0	7 927	3 637	1 0310	6 886
1	005	850	005	300	25	6 849	6 779	1 937	1318 0	7 916	3 638	1 0295	6 888
2	05	850	01	300	25	6 858	6 744	1 929	1310 0	7 912	3 644	1 0290	6 899
1	1 00	-----	01	575	1 00	6 123	6 100	1 967	166 7	7 896	3 997	1 0269	7 567
1	1 00	-----	005	575	1 00	6 054	6 069	1 947	150 7	7 893	4 005	1 0265	7 582
1	002	800	005	500	75	5 921	6 089	1 936	810 7	7 839	4 029	1 0195	7 628
1	1 00	-----	05	650	1 00	5 863	5 559	1 882	724 8	7 792	4 802	1 0134	9 091
1	1 00	-----	025	650	1 00	5 357	5 285	1 835	454 0	7 765	4 882	1 0099	9 243
1	10	600	025	650	1 00	5 357	5 285	1 835	454 0	7 765	4 882	1 0099	9 243
2	1 00	-----	025	650	1 00	5 221	5 096	1 825	264 1	7 761	4 932	1 0094	9 337
2	10	600	025	650	1 00	5 221	5 096	1 825	264 1	7 761	4 932	1 0094	9 337
1	1 00	-----	01	600	1 00	5 157	5 113	1 824	275 6	7 758	4 933	1 0090	9 339
1	1 00	-----	01	650	1 00	5 139	5 113	1 825	275 6	7 758	4 933	1 0090	9 339
1	10	600	01	650	1 00	5 139	5 113	1 825	275 6	7 758	4 933	1 0090	9 339
1	01	700	01	600	1 00	5 157	5 113	1 824	275 6	7 758	4 933	1 0090	9 339
1	1 00	-----	005	600	1 00	5 073	5 078	1 817	238 2	7 750	4 942	1 0080	9 356
1	01	700	005	600	1 00	5 073	5 078	1 817	238.2	7 750	4 942	1 0080	9 356
1	1 00	-----	01	675	1 00	4 170	4 124	2 015	454 7	7 740	5 999	1 0066	11 357
2	1 00	-----	025	725	1 00	4 243	4 104	2 012	430 4	7 730	6 000	1 0053	11 359
1	005	750	01	675	1 00	4 084	4 086	2 025	372 4	7 716	6 011	1 0035	11 380
1	005	850	01	675	1 00	4 084	4 086	2 025	372 4	7 716	6 011	1 0035	11 380
3	25	500	025	725	1 00	4 179	4 041	1 996	288 0	7 709	6 023	1 0026	11 403
3	1 00	-----	025	725	1 00	4 152	4 032	1 996	268 4	7 708	6 026	1 0025	11 409
3	25	700	025	725	1 00	4 152	4 032	1 996	268 4	7 708	6 026	1 0025	11 409
4	1 00	-----	05	675	1 00	4 304	4 008	1 987	206 5	7 706	6 035	1 0022	11 426
5	1 00	-----	05	675	1 00	4 225	4 000	1 973	185 4	7 689	6 038	1 0000	11 431

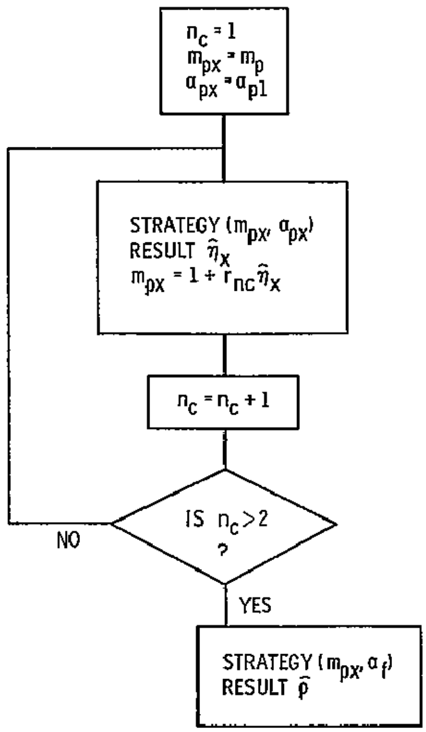
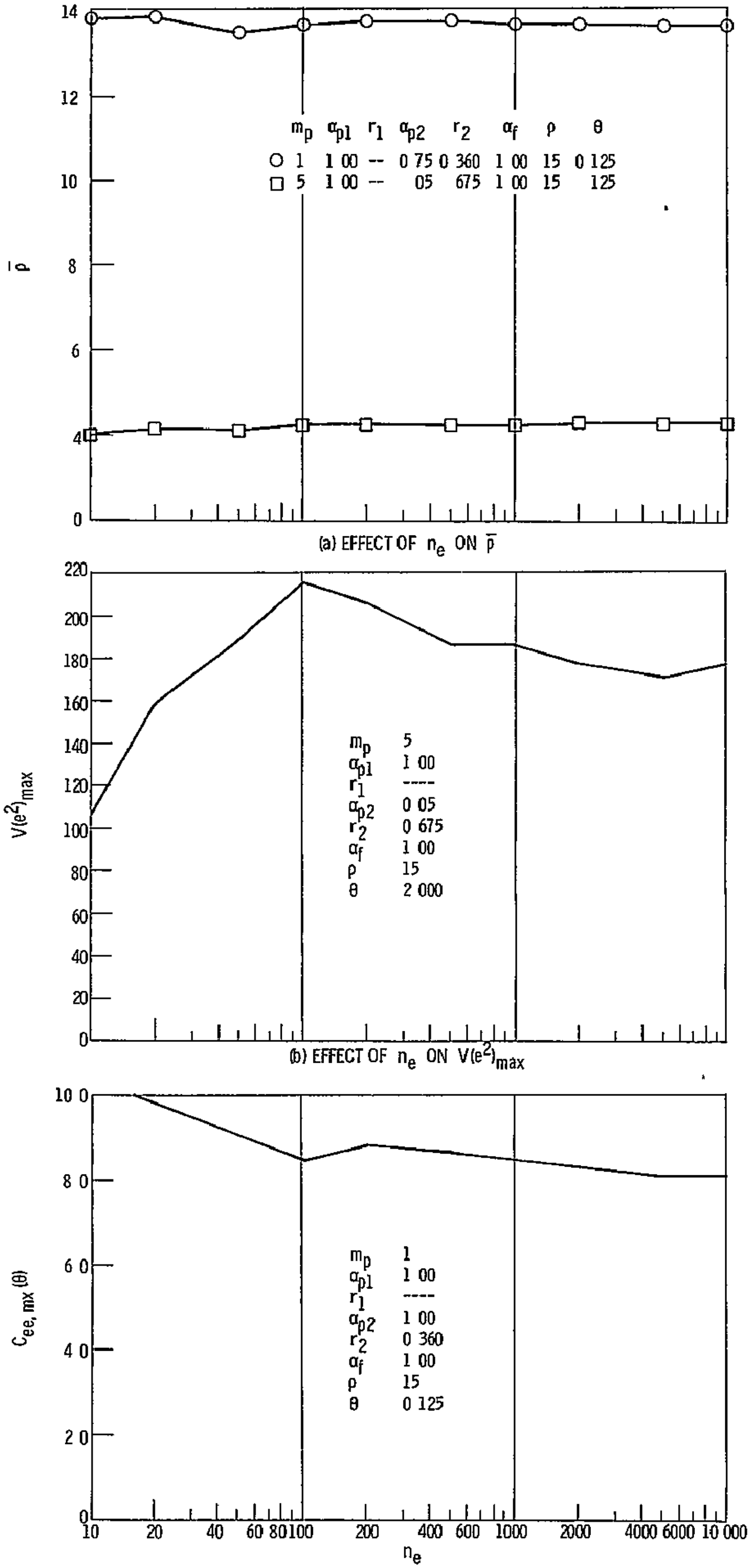


Figure 1 – Strategy (m_p, α_{p1}, r_1, α_{p2}, r_2, α_f)



(a) EFFECT OF n_e ON $\bar{p}$

(b) EFFECT OF n_e ON $V(e^2)_{max}$

(c) EFFECT OF n_e ON $C_{ee\ mx}(\theta)$

Figure 2 - Effect of sample size on Monte Carlo results

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON D C 20546

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

SPECIAL FOURTH-CLASS RATE
BOOK

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
451

U.S.MAIL

POSTMASTER If Undeliverable (Section 158 Postal Manual) Do Not Return