

Department of Statistics

UNIVERSITY OF WISCONSIN

Madison, Wisconsin

Technical Report No. 2

August 1961

A USEFUL METHOD FOR MODEL-BUILDING

by

G. E. P. Box and William Hunter

This research was supported by the United States Navy through the Office of Naval Research, under Contract Nonr-1202(17), Project NR 042 222. Reproduction in whole or in part is permitted for any purpose of the United States Government.

A USEFUL METHOD FOR MODEL BUILDING

by

G. E.P. Box and William G. Hunter

The object of much experimentation is to build or discover a suitable model. This is done by an iterative procedure in which a particular model is tentatively entertained, strained in various ways over the region of application, and its defects found. The nature of the defects interacting with the experimenter's technical knowledge can suggest changes and remedies leading to a new model which, in turn, is tentatively entertained and submitted to a similar straining process. In this paper a simple method is presented for checking and modifying a model in which the estimated constants of the system are treated as the observations. We have no reason to suppose that this method is new but we feel that it is of sufficient general interest to be formally recorded.

1. Empirical vs. Theoretical Models

In some cases so little is known about a physical, chemical, or biological process that in studying it we are forced to adopt a heavily empirical approach which assumes little more than the process function can be represented as locally smooth. In these circumstances

we may be able to do no more than represent the relation by a graduating function such as a polynomial. Such a graduating function acts as a mathematical French curve and makes no claim of course to represent functionally the true relationship. At the other extreme, cases may occur in which a theoretical model based on an intimate knowledge of the physical mechanism is available. In such a case the functional relation is often defined by time- or space-dependent differential equations. If a model adequately represents a given situation, we can estimate the behavior of the system to a known degree of precision by estimating the constants and their precision [1,2,4,7]. In the first case we know practically nothing about the functional form; in the second, we know everything about the functional form.

Most frequently, the situation is somewhere in between the two extremes. The theory is usually inadequately understood, but by a series of planned experiments using an iterative process like that already described we hope to arrive at a working model which takes account of the main theoretical aspects of the problem. Strictly speaking, no model can ever give a precise description of what happens. A working theoretical model, however, supplies information on the system under study over important ranges

of the variables by means of equations which reflect at least the major features of the mechanism. Quite possibly it fails to hold over more extensive ranges where effects which previously could be safely ignored become important.

It is common to find too timid an attitude toward model-building. There is a tendency to think that unless the functional form is almost exactly known, the theoretical approach is unjustified. This results in many purely empirical studies where a semi-theoretical approach in which knowledge of the functional form was built up step by step would have been more effective and more rewarding.

The functional form derived from a theoretical development is usually non-linear in the unknown parameters. The empirical approach is sometimes adopted merely to avoid the complication of fitting these non-linear equations. The fitting of non-linear equations is chiefly a task in numerical analysis, and various devices can be used to make it reasonably simple. In particular the present availability of electronic computers has largely removed the computational labor which would otherwise make this approach unattractive.

The effect of adopting the empirical approach, when a theoretical one would have been possible, is often to indirectly invite trouble. In a purely empirical formulation non-linearities in the form of interactions and complex

effects may arise which can be largely avoided by a semi-theoretical approach. Not only are such non-linearities difficult to interpret but it may be necessary to employ an excessively large number of experiments to estimate the proliferation of constants which result from the use of an inappropriate empirical functional form.

In this report we show how by working with the estimated parameters of the problem and treating these rather than the original responses as the "observations," this introduction of unnecessary non-linearity is avoided, often-times with great simplification in the experimentation required. The method can best be illustrated by means of an example.

2. A Chemical Example with Four Variables

In Table 1 are shown the results which might have been obtained from a 2^4 factorial experiment on a chemical process in which the variables studied were $[A]_0$ and $[B]_0$ the initial concentrations of reactants A and B, the concentration $[C]$ of catalyst C, and the temperature D. No confusion will arise if we denote these variables in the factorial experiment by A, B, C, and D. During each run, samples were taken at five times (80, 160, 320, 640 and 1280 minutes), and the yield (the concentration of the product F) was determined.

Table 1 Experimental Results

Run	Design A B C D	Concentration of F (moles per liter $\times 10^2$)				
		t=80 min	160 min	320 min	640 min	1280 min
1	- - - -	3.17	5.39	8.66	15.9	22.6
2	+ - - -	14.7	23.4	34.3	34.6	20.3
3	- + - -	4.80	10.8	22.5	34.6	42.0
4	+ + - -	23.2	39.0	55.6	63.4	41.6
5	- - + -	3.72	3.81	17.2	20.0	23.9
6	+ - + -	17.9	28.3	40.5	34.2	21.6
7	- + + -	8.60	13.3	25.9	39.8	50.8
8	+ + + -	30.9	51.4	72.2	76.4	38.9
9	- - - +	7.48	9.93	20.0	30.9	24.9
10	+ - - +	25.3	35.3	39.1	28.4	7.50
11	- + - +	13.3	27.1	43.0	58.0	49.4
12	+ + - +	50.8	75.6	84.2	57.0	11.5
13	- - + +	9.15	15.8	27.5	33.9	23.0
14	+ - + +	30.8	44.4	46.7	24.9	2.94
15	- + + +	22.8	37.2	57.9	69.1	53.9
16	+ + + +	62.6	88.0	89.5	43.4	5.80

	A (moles/liter)	B(moles/liter)	C(millimoles/liter)	D($^{\circ}$ C)
+	40	2	1.0	175
-	20	1	0.5	165

2.1 Conventional Analysis

The usual factorial analysis of these results is shown in Table 2. We suppose that the between-run standard deviation in the units of concentration is known to be of the order of 2 so that the standard deviation associated with the effects in Table 2 is about 1. We therefore appear to be dealing with an exceedingly complex phenomenon in which two-factor interactions and three-factor interactions as well as main effects are of importance. Furthermore, the main effects and interactions change markedly and in a complex fashion with regard to time. (Note, for example, the A effect, the AD and the ABD interactions.) If we regarded time as a further factor in a "split-plot" design we would be faced with the interpretation of a five-factor interaction.

Table 2. Factorial Analysis of the Yields

Effects (moles per liter $\times 10^2$)	t(min)	80	160	320	640	1280
A		21.45	30.32	26.46	3.55	-20.15
B		11.65	19.57	23.63	23.40	15.80
C		6.92	9.40	12.23	6.33	2.73
D		15.86	22.18	19.86	7.30	- 7.74
AB		5.15	6.20	4.65	- 1.80	- 9.63
AC		3.04	5.18	3.65	0.48	- 0.45
BC		4.18	4.83	4.77	5.52	3.70
AD		7.74	8.00	1.32	-13.10	+10.72
BD		7.54	11.05	11.70	3.95	- 0.23
CD		2.02	- 0.31	- 3.40	- 7.08	- 4.64
ABC		1.42	2.75	4.20	3.22	0.22
ABD		4.30	5.13	3.98	- 2.00	- 2.50
ACD		- 1.50	- 3.79	- 6.02	- 8.28	- 2.77
BCD		- 0.65	- 2.94	- 3.50	- 6.02	- 2.38
ABCD		- 1.80	- 2.98	- 6.63	- 7.78	- 2.10
MEAN		19.85	30.58	41.06	39.55	26.24

2.2 Alternative Approach

Fortunately, in this investigation the experimenter was able to conceive of a theoretical model which might act at least as a starting point in explaining the time dependency of the data. The reaction studied was of the type



where F was the desired product. The concentration of $[A]_0$ employed was large compared with that of $[B]_0$, roughly in the ratio of 20 to 1. Thus, the percentage of A used up during the reaction was small and consequently, as an approximation, its concentration could be treated as constant for any given run.

Tentative Model

The simplest possible kinetic model describing the behaviour of the system would be one which assumed that the reactions were first order with respect to the concentrations of B and F. Then for a given initial concentration of the reactant A, a given concentration of the catalyst C and a given temperature D, the course of the reaction would be described by the following set of differential equations.

$$- \frac{d[A]}{dt} = k_1[B] + k_2[F]$$

$$- \frac{d[B]}{dt} = k_1[B]$$

$$\frac{d[F]}{dt} = k_1[B] - k_2[F]$$

$$\frac{d[G]}{dt} = k_2[F]$$

where $[B]$, $[F]$, $[G]$ refer to concentrations at time t .

The rate equations can be integrated to give for the concentration of F ,

$$[F] = \frac{[B]_0 k_1}{k_1 - k_2} (e^{-k_2 t} - e^{-k_1 t}) \quad (1)$$

In this expression $k_1 = k_1([A]_0, [C], D)$ and $k_2 = k_2([A]_0, [C], D)$ would be expected to be functions of the experimental variables $[A]_0$, $[C]$, and D but not of $[B]_0$ if the model is adequate.

For a given experimental run we have five observations of the concentration of F at different times (Table 1) and from these we can estimate k_1 and k_2 by the method of least squares. To avoid digression from the main theme at this stage, we discuss this estimation procedure in the Appendix. In this appendix we also illustrate the use of the likelihood surface in the appreciation of the estimation situation.

The fitted curves for each of the 16 runs are shown in Figure 1. Careful inspection of the data points in relation to these curves and of the residuals shown in Table 4 gives no reason to doubt the adequacy of the assumptions so far.

Analysis of the Rate "Constants"

There is some advantage in working with the logarithms rather than the rate "constants" themselves. In Table 3 we show $-10 \ln \hat{k}_1$ and $-10 \ln \hat{k}_2$ for each of the 16 runs. The factorial analysis applied to these values is shown in Table 5.

We notice first that if our assumptions were true that the reactions were first order with respect to the concentration of reactant B, then no main effects or interactions containing B should be present. Inspection of the effects in Table 5 provides no reason to doubt this assumption.

We have now to consider the probable nature of the dependence of k_1 and k_2 on the remaining three factors A, C, and D - the concentration of A, the catalyst concentration and the temperature. In accordance with simple kinetic theory, it might be expected that the temperature dependence of the rate constants will be given by the Arrhenius law, that is to say that $\ln k_i$ ($i = 1, 2$) will

Table 3. Least Squares Estimates of "Constants"

Run	$-10 \ln \hat{k}_1$	$-10 \ln \hat{k}_2$
1	79.80	72.56
2	61.99	62.09
3	77.49	68.03
4	64.56	62.04
5	74.80	67.25
6	59.91	62.19
7	75.99	69.35
8	61.01	62.40
9	69.48	65.09
10	55.70	58.05
11	68.78	64.31
12	54.82	58.19
13	65.71	63.42
14	52.04	57.86
15	65.22	64.44
16	51.76	57.01

Table 4. Residuals

$$[y_t - \eta_t(\hat{k}_{1j}, \hat{k}_{2j})] \times 100 \quad \left\{ \begin{array}{l} t = 80, \dots, 1280 \\ j = 1, 2, \dots, 16 \end{array} \right\}$$

Run	t = 80 min	160 min	320 min	640 min	1280 min
1	0.55	0.35	-0.60	0.22	0.01
2	0.90	-0.08	0.31	-1.01	0.76
3	-1.68	-1.38	0.84	0.72	-0.36
4	1.38	1.20	-1.12	-0.58	0.70
5	-0.49	-4.03	3.57	-0.66	-0.07
6	1.21	0.40	1.47	-4.04	3.13
7	1.04	-0.96	0.64	-0.38	0.10
8	0.48	0.12	-1.18	1.26	-0.68
9	0.52	-2.70	-0.78	2.71	-1.22
10	2.12	-0.01	-1.91	0.64	1.04
11	-1.58	0.42	-0.40	0.88	-0.44
12	0.78	0.28	-1.40	1.22	-0.76
13	-0.72	-1.60	0.48	1.27	-0.87
14	-0.45	-0.18	0.92	-0.12	-1.31
15	1.92	0.28	0.22	-1.56	0.92
16	-0.56	-0.56	1.84	-1.12	-0.64

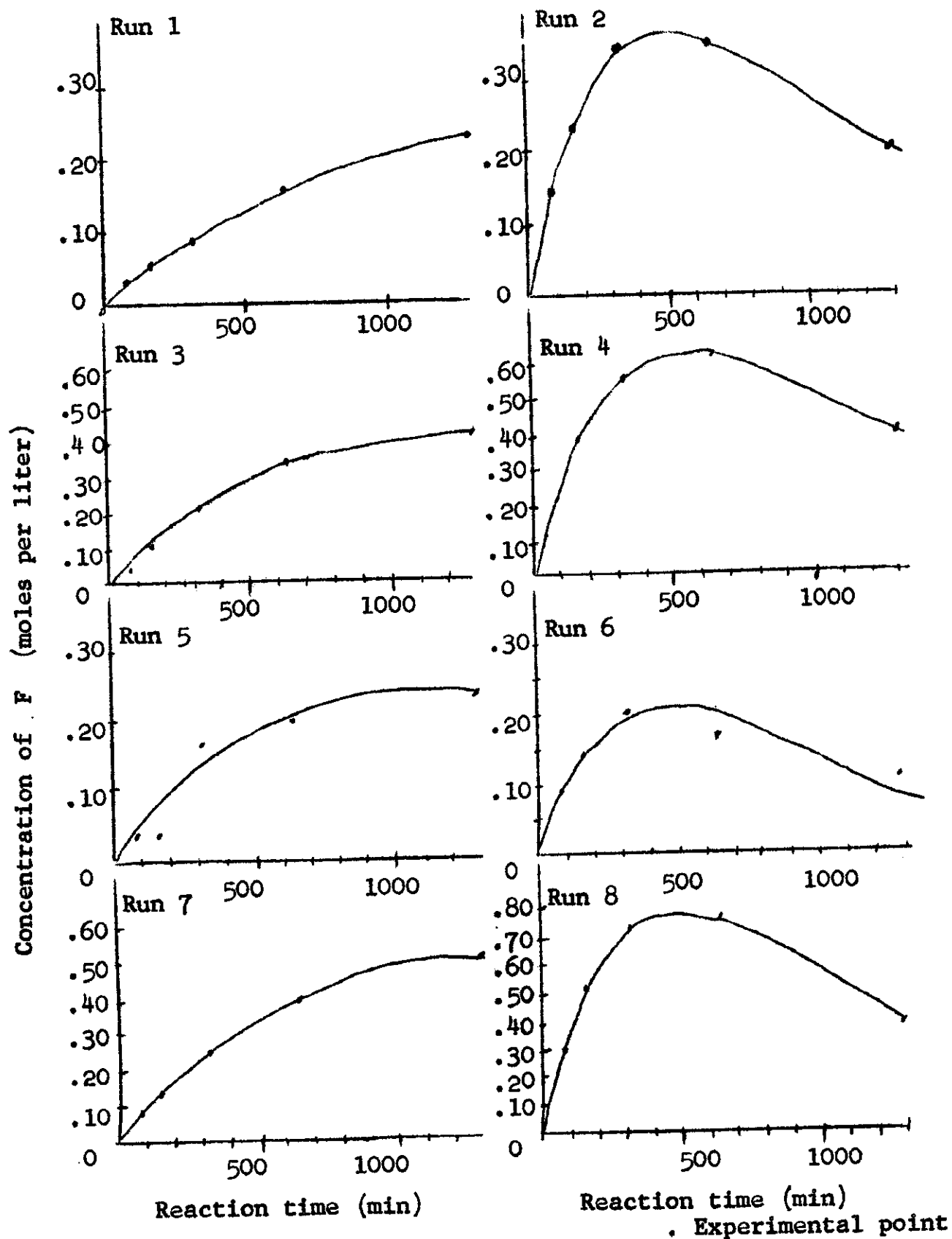
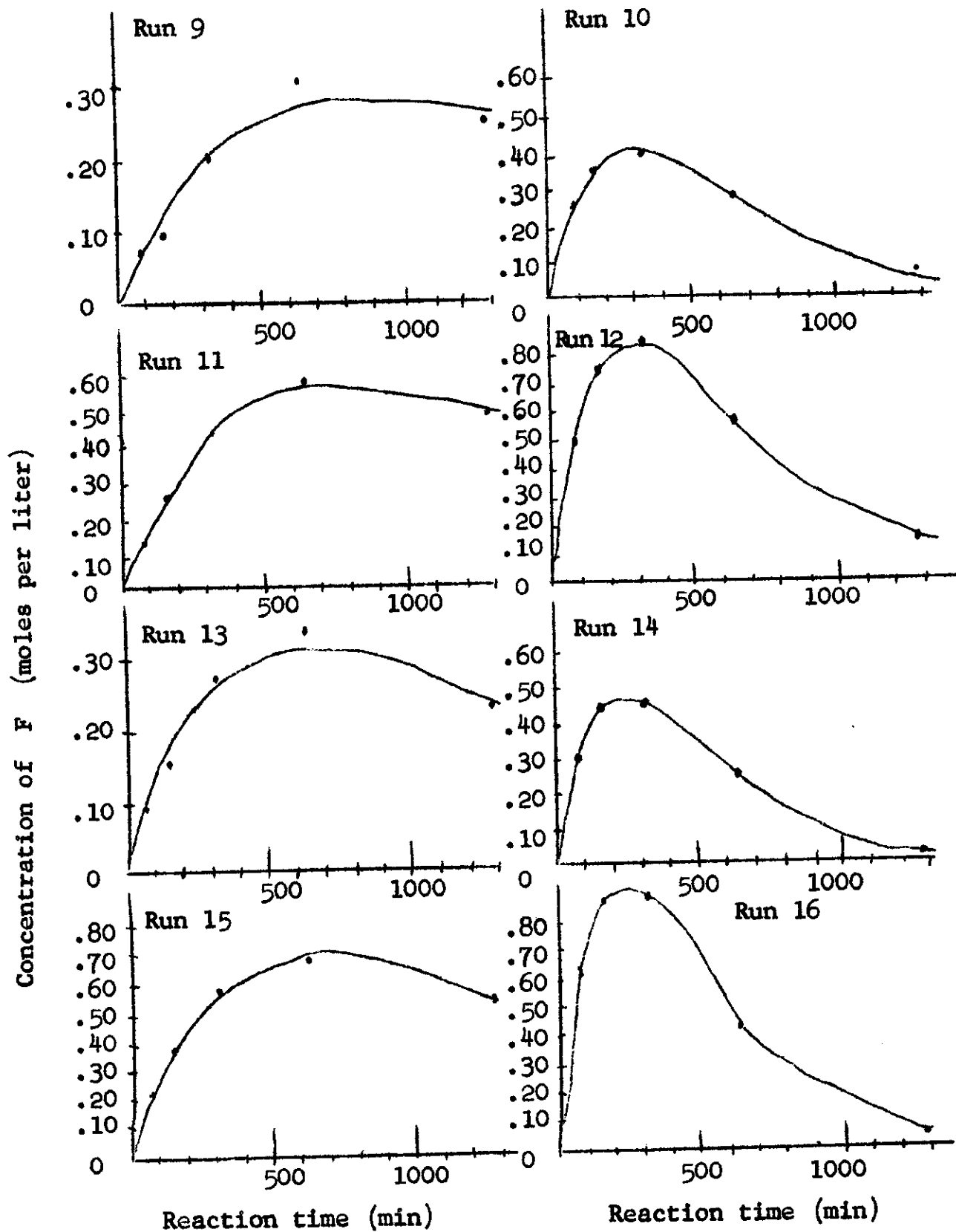


Figure 1. Fitted Curves



. Experimental point

Figure 1. Fitted Curves

Table 5. Factorial Analysis of $\ln \hat{k}_1$ and $\ln \hat{k}_2$

Effects	on $10 \ln \hat{k}_1$	on $10 \ln \hat{k}_2$
A	14.44	6.83
B	-0.02	0.34
C	3.27	0.80
D	9.00	4.69
AB	-0.60	-0.20
AC	-0.18	-0.58
BC	-0.36	-0.96
AD	-0.72	-0.29
BD	0.61	-0.22
CD	0.24	-0.08
ABC	0.57	1.14
ABD	0.60	0.44
ACD	0.03	0.54
BCD	0.15	0.76
ABCD	-0.67	-0.45
MEAN	-64.94	-63.39

be a linear function of the reciprocal absolute temperature. Also, experimental evidence supported at least in part by kinetic theory suggests that the constants will be proportional to some power of the concentration of A and to some power of the catalyst concentration. Thus, the equations

$$k_1 = k_1([A]_0, [C], T) = [A]_0^{p_1} [C]^{q_1} \alpha_1 e^{-\beta_1/T}$$

$$k_2 = k_2([A]_0, [C], T) = [A]_0^{p_2} [C]^{q_2} \alpha_2 e^{-\beta_2/T}$$

are now tentatively entertained as describing the dependence of the rate constants on the variables. Equivalently,

$$\ln k_1 = p_1 \ln[A]_0 + q_1 \ln[C] + \ln \alpha_1 - \beta_1/T \quad (2)$$

$$\ln k_2 = p_2 \ln[A]_0 + q_2 \ln[C] + \ln \alpha_2 - \beta_2/T$$

Since these expressions are linear in $\ln[A]_0$, $\ln[C]$, and T^{-1} , all interactions involving $[A]_0$, $[C]$, and T would be zero if our tentative assumptions were correct. Again, consideration of Table 5 lends support to our hypotheses. Inspection of these interactions and plotting on probability paper [6] show no abnormalities.

We shall proceed, therefore, on the assumption that Equations (2) provide an adequate representation of the

dependence of the constants on the factors A , C , and D .

With this same assumption, the pooled interactions supply an estimate of the standard errors of the effects.* Denoting by s_1 and s_2 the standard errors of the effects calculated from $\ln \hat{k}_1$ and $\ln \hat{k}_2$ respectively we have, for example,

$$s_1^2 = \frac{(-.060)^2 + (-.018)^2 + \dots + (-.067)^2}{11} = 0.00239 , \text{ or}$$

$$s_1 = 0.049$$

In a similar way we have $s_2 = 0.069$. The 95% confidence intervals for the effects calculated from $\ln \hat{k}_1$ and $\ln \hat{k}_2$ respectively are therefore given by

Estimated effect from $\ln \hat{k}_1 \pm 2.201 \times 0.049$

Estimated effect from $\ln \hat{k}_2 \pm 2.201 \times 0.060$

where 2.201 is the 5% level of Student's t with eleven degrees of freedom.

Having checked our assumptions, it now remains to estimate the unknowns in Equations (2) . Denoting averages by bars, we have for the "average log rate constants"

* One would not expect in this example that the variances of the estimated constants would necessarily remain the same from run to run. It may be shown, however, that for the case of a two-level factorial design, which is here employed, unbiased estimates of the effects and their standard errors are obtained even so [3] .

$$\overline{\ln k_i} = p_i \overline{\ln[A]_O} + q_i \overline{\ln[C]} + \ln \alpha_i - \beta_i \frac{1}{T} \quad (i = 1, 2)$$

On subtraction from Equations (2) we obtain

$$\ln k_1 = \overline{\ln k_1} + p_1(\ln[A]_O - \overline{\ln[A]_O}) + q_1(\ln[C] - \overline{\ln[C]}) - \beta_1\left(\frac{1}{T} - \frac{1}{\overline{T}}\right) \quad (3)$$

$$\ln k_2 = \overline{\ln k_2} + p_2(\ln[A]_O - \overline{\ln[A]_O}) + q_2(\ln[C] - \overline{\ln[C]}) - \beta_2\left(\frac{1}{T} - \frac{1}{\overline{T}}\right)$$

Estimates of the unknowns are thus provided by

$$\hat{p}_i = \frac{A_i \text{ effect}}{\ln[A]_O^+ - \ln[A]_O^-}, \quad \hat{q}_i = \frac{C_i \text{ effect}}{\ln[C]^+ - \ln[C]^-}, \quad \hat{\beta}_i = \frac{D_i \text{ effect}}{\frac{1}{T^-} - \frac{1}{T^+}} \quad (i=1,2)$$

where the A_1 effect refers to the "A effect" calculated from $\ln \hat{k}_1$, the " A_2 effect" that calculated from $\ln \hat{k}_2$.

The 95% confidence intervals for these unknowns are

$$\hat{p}_i \pm \frac{2.201 s_i}{\ln[A]_O^+ - \ln[A]_O^-}$$

$$\hat{q}_i \pm \frac{2.201 s_i}{\ln[C]^+ - \ln[C]^-}$$

$$\hat{\beta}_i \pm \frac{2.201 s_i}{\frac{1}{T^-} - \frac{1}{T^+}}$$

Thus finally we obtain

$$\hat{p}_1 = 2.08 \pm 0.16$$

$$\hat{p}_2 = 0.99 \pm 0.20$$

$$\hat{q}_1 = 0.47 \pm 0.16$$

$$\hat{q}_2 = 0.12 \pm 0.20$$

$$\hat{\beta}_1 = (17.7 \pm 2.1) \times 10^3$$

$$\hat{\beta}_2 = (9.2 \pm 2.7) \times 10^3$$

$$\ln \alpha_1 = 26.66 \pm 0.52$$

$$\ln \alpha_2 = 11.16 \pm 0.65$$

The activation energies E of the two reactions are given by

$$E = R \beta$$

where R is the gas constant $1.987 \text{ cal. deg.}^{-1} \text{ mole}^{-1}$.

Thus
$$E_1 = 35.2 \pm 4.2 \text{ kcal}$$

$$E_2 = 18.3 \pm 5.4 \text{ kcal.}$$

Kinetic theory renders most plausible the existence of reactions of orders zero, one, two, or three for the concentrations of A and B. The situation with regard to catalyst concentration is less well understood; however, in some circumstances half-order reactions are plausible on theoretical grounds.

We see that the data are readily explained in terms of a kinetic model which behaves as if the first reaction ($A + B \rightarrow F$) is second-order with respect to A, first-order with respect to B, and half-order with respect to the catalyst concentration. The situation with regard to the second reaction ($A + F \rightarrow G$) is less well interpreted because of the larger errors of estimate arising with these constants; however, we may say that the second reaction is first order with respect to [A] and [B] while the effect of catalyst concentration appears to be slight. Actual trial shows that the function

$$[F] = \frac{[B]_0 [A]_0^2 [C]^{0.5} k_1'}{[A]_0^2 [C]^{0.5} k_1' - [A]_0 k_2'} \left(e^{-[A]_0 k_2' t} - e^{-[A]_0^2 [C]^{0.5} k_1' t} \right)$$

where

$$k_1' = 3.786 \times 10^{11} e^{-35200/RT}$$

$$k_2' = 7.026 \times 10^4 e^{-18300/RT}$$

closely reproduces the yields of F at the various times and the various experimental conditions. We can with some confidence use it for prediction purposes at least over the range of variables actually tested.

3. Discussion

The device adopted is of rather general application. Many problems in chemistry, engineering, physics, and biology appear to be of a kind which could usefully be tackled in this manner. In those instances where the functional form is defined by differential equations which are not integrable to explicit solutions, numerical integration may be employed. The Princeton-IBM 704 program [4] is one example of a general computer program which uses a suitable subroutine for numerically solving the equations and automatically applies an iterative procedure to obtain the least squares solution. The analog computer provides an alternative means for fitting the constants although care should be taken when using these machines to explore the estimation situation by obtaining at least a rough plot of the sum of squares surface. If this is not done, excessive ill-conditioning of the sum of squares surface may go unnoticed and lead to serious trouble.

The advantage of an analysis of the "constants" of the system is that unnecessary non-linearities are removed and inadequacies in the model can be shown up in readily understood form. Where these inadequacies are of minor character and not such as to cast doubt

on the general usefulness of the basic model, allowances are readily made. We are often in one of two simple situations. Either (1) the effects of the variables have already been taken into account in the model so that no effects will occur (this is the case in the example with respect to the initial concentration of reactant B); or, alternatively, (2) the only effects that occur will be linear (this is the case for the variables A , C and D).

Our aim, therefore, should be to subject our model to the largest possible number of "strains" by varying each of the variables which should theoretically have no effect or possibly only a linear effect. The existence of non-zero effects or interactions indicates the necessity for modification . It is seen that we are here in an ideal situation for the application of fractional factorial designs. Such designs are particularly valuable where experiments are expensive and the number of runs necessarily limited. When maximum economy is essential, fractional factorial designs which are of particular value are those of resolution IV [5]. For example using one such design, the behavior of the model with respect to as many as three principal variables and eight minor variables can be checked in only 16 runs. The principal variables are those which it is feared might give trouble.

by showing not only main effects but also interactions; and the minor variables are those to be checked only in so far as their main effects are concerned. In this context in the above example the minor variables would include the initial concentration of B and any other variables the effects of which it is believed are already taken into account in the model and which would, therefore, be expected to have zero effect.

In our example we have illustrated the method using a factorial design which supplies a very efficient way of exerting appropriate strains on the model from which the data are readily analyzed. It will be noticed that this same method can be applied, although less efficiently when the levels of the variables do not follow a prescribed design. The constants will then appear as the dependent variables in a multiple regression equation. This particular use of multiple regression is almost unique since it is one case where the assumption of the linear model may reasonably be expected to be justified.

A P P E N D I X

Least Squares Estimation of k_1 and k_2

The constants are found by an iterative least squares method using an approximate linearization of the function (1). It is often possible to obtain a closer linear approximation by expanding in terms of suitable functions of the parameters rather than the untransformed variables themselves [4]. For the particular functional form here considered it is readily shown by considering higher order terms, that expansion in terms of $\ln k_1$ and $\ln k_2$ provides a closer approximation than the more direct expansion in k_1 and k_2 .

If we let $\theta_1 = 4 + \log k_1$ and $\theta_2 = 4 + \log k_2$ and denote the function by $\eta(\theta_1, \theta_2)$, we have

$$\eta(\theta_1, \theta_2) = \eta(\theta_1^*, \theta_2^*) + (\theta_1 - \theta_1^*) \left[\frac{\partial \eta}{\partial \theta_1} \right]_{\theta = \theta^*} + (\theta_2 - \theta_2^*) \left[\frac{\partial \eta}{\partial \theta_2} \right]_{\theta = \theta^*} + \epsilon \quad (3)$$

ignoring terms of degree higher than the first in $(\theta - \theta^*)$. In the common case in which the tentative function is not known explicitly but is for example given only in terms of differential equations which can be solved numerically, the derivative can be replaced by a difference approximation. The difference method is usually most convenient even when direct differentiation is available and we shall use it here. Letting θ_1^* denote our first guess of the true value $\theta_1 (i = 1, 2)$,

$$x_1 = \eta(\theta_1^* + \delta_1, \theta_2^*) - \eta(\theta_1^*, \theta_2^*) \quad x_2 = \eta(\theta_1^*, \theta_2^* + \delta_2) - \eta(\theta_1^*, \theta_2^*) \quad (4)$$

$$\delta_1 = \theta_1 - \theta_1^* \quad \delta_2 = \theta_2 - \theta_2^* \quad (5)$$

$$Y = y - \eta(\theta_1^*, \theta_2^*) \quad (6)$$

Equation (3) takes on the simple "regression" form

$$Y = \beta_1 x_1 + \beta_2 x_2 + \epsilon . \quad (7)$$

Estimates b_1 and b_2 for β_1 and β_2 can be obtained by the usual method of least squares and may be used to obtain new estimates θ_1^{**} and θ_2^{**} of θ_1 and θ_2 using

$$\begin{aligned} \theta_1^{**} &= \theta_1^* + \delta_1 b_1 \\ \theta_2^{**} &= \theta_2^* + \delta_2 b_2 . \end{aligned} \quad (8)$$

Provided the iteration converges, these new estimates will be closer to the least squares values and may in turn be used as starting values for further iteration. The process can be repeated until there is no material improvement.

For our present purpose we shall show the estimation of k_1 and k_2 for the first run. The rapidity of convergence of this non-linear estimation method depends markedly on the closeness of the original guesses k_1^* and k_2^* . It is always worthwhile to take a little trouble in obtaining reasonably good preliminary estimates.

For the present example, the preliminary estimates

were found as follows. Using Equation (1) we have at time zero

$$\left(\frac{d[F]}{dt}\right)_{t=0} = k_1[B]_0$$

whence

$$k_1 = \left(\frac{d[F]}{dt}\right)_{t=0} / [B]_0 \quad (9)$$

Also if t_{\max} represents the time at which $[F]$ attains its maximum value, then it is readily shown by direct differentiation that

$$t_{\max} = \frac{\ln k_1 - \ln k_2}{k_1 - k_2} \quad (10)$$

We can therefore proceed by making a rough sketch of the curve of the yield of F against time for a particular run, estimating $(d[F]/dt)_{t=0}$ and t_{\max} , and substituting these estimates in Equations (9) and (10) to obtain starting values k_1^* and k_2^* for k_1 and k_2 . For illustration, a plot of the curve for the first run is shown in Figure 2.

Figure 2.

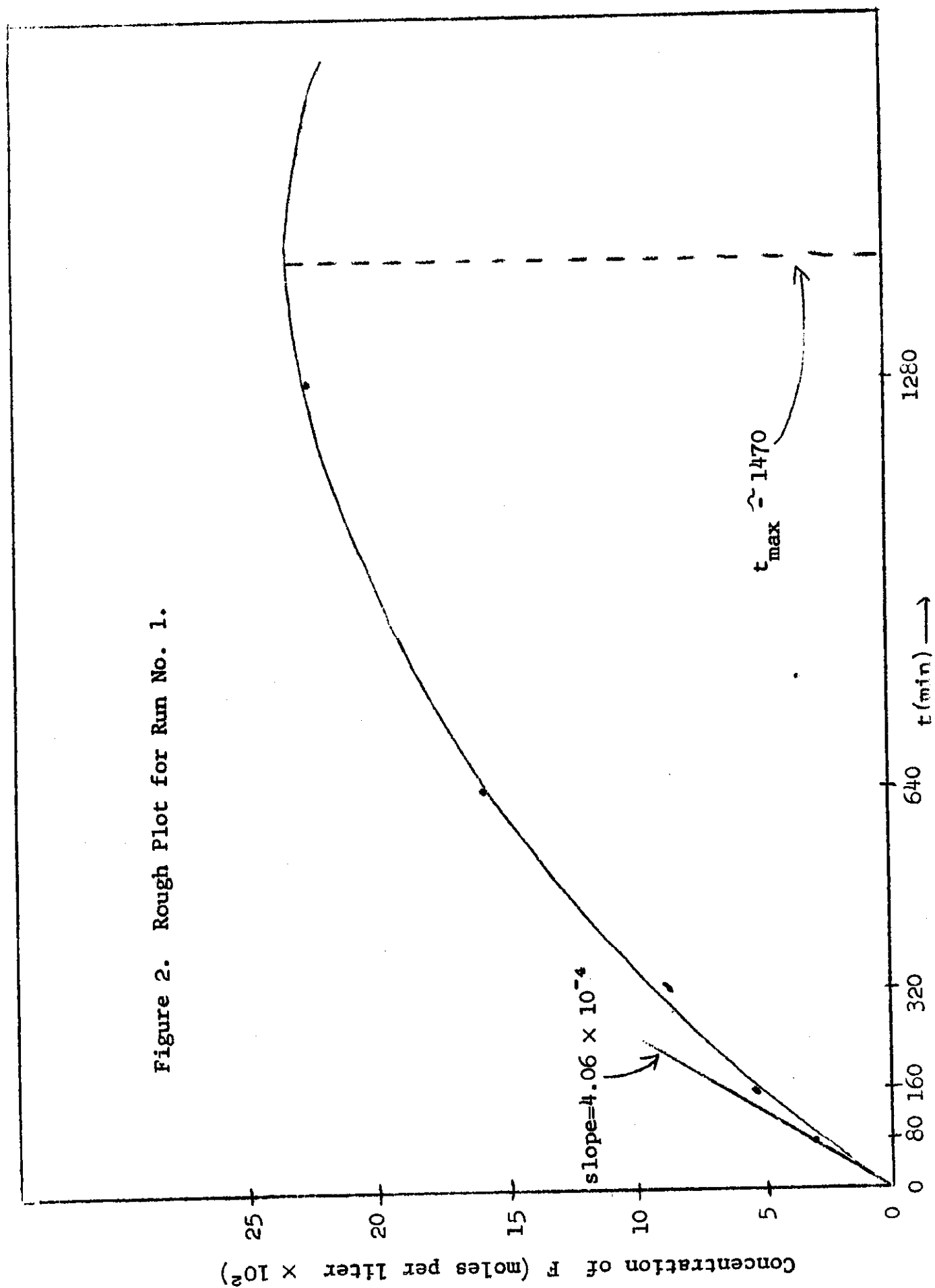
The values read off from this curve are

$$(dF/dt)_{t=0} = 4.06 \times 10^{-4} \text{ moles per liter/min. and } t_{\max} = 1470 \text{ min.}$$

yielding the preliminary estimates $k_1^* = 4.06 \times 10^{-4} \text{ min}^{-1}$

and $k_2^* = 10.52 \times 10^{-4} \text{ min}^{-1}$. Using these estimates

Figure 2. Rough Plot for Run No. 1.



as starting values we can now carry through the more precise estimation of k_1 and k_2 by least squares.

For the present example we have $\theta_1^* = 0.6090$
and $\theta_2^* = 1.0220$. Taking $\delta_1 = .05$ and $\delta_2 = .02$ we readily obtain

	{1}	{2}	{3}	{4}
t(min)	observed value y	$\eta(0.61, 1.02)$	$\eta(0.66, 1.02)$	$\eta(0.61, 1.04)$
80	.0317	.0307	.0345	.0306
160	.0539	.0579	.0648	.0577
320	.0866	.1032	.1149	.1024
640	.1590	.1643	.1815	.1618
1280	.2260	.2104	.2281	.2047

Thus	{1} - {2}	{3} - {2}	{4} - {2}
	Y	x_1	x_2
	.0010	.0038	-.0001
	-.0040	.0069	-.0002
	-.0166	.0117	-.0008
	-.0053	.0172	-.0025
	.0156	.0177	-.0057

$$\text{And } \Sigma x_1^2 = 80.807 \times 10^{-5} \quad \Sigma x_2^2 = 3.943 \times 10^{-5} \quad \Sigma x_1 x_2 = -15.501 \times 10^{-5}$$

$$\Sigma Y x_1 = -3.306 \times 10^{-5} \quad \Sigma Y x_2 = -6.169 \times 10^{-5}$$

In this case we obtain the Normal equations

$$80.807 b_1 - 15.501 b_2 = -3.306$$

$$-15.501 b_1 + 3.943 b_2 = -6.169$$

whence $b_1 = -1.388$ and $b_2 = -7.021$

$$\text{and } \theta_1^{**} = 0.6090 + (.05)(-1.388) = 0.5396$$

$$\theta_2^{**} = 1.0220 + (.02)(-7.021) = 0.8816$$

These new estimates θ_1^{**} and θ_2^{**} are then used as the starting values for further iteration. The course of the iteration up to the third stage is shown below.

Iteration	θ_1	θ_2
1 st	0.6090	1.0220
2 nd	0.5396	0.8816
3 rd	0.5340	0.8485

So we take as the final estimates $k_1 = 3.42 \times 10^{-4} \text{ min.}^{-1}$

$k_2 = 7.06 \times 10^{-4} \text{ min.}^{-1}$. A similar process can be used

to find estimates of the constants for each of the sixteen runs. The logarithms of these values are shown in Table 3.

When, as in this example, the function is not markedly non-linear in the parameters over the region in which the iteration takes place, it may be unnecessary to calculate $\sum x_1^2$, $\sum x_2^2$, and $\sum x_1 x_2$ at each iteration. In the present example, for instance, it is found that these quantities remain essentially the same from one trial to the next. In general if hand calculation is employed, considerable saving is

achieved when these values can be reused from one iteration to the next.

Before proceeding further with the analysis, it is of course essential to confirm that the assumptions underlying the calculation of these preliminary estimates are reasonable. In a non-linear situation it is necessary to make a careful and critical study of the estimation situation. Questions requiring answers are:

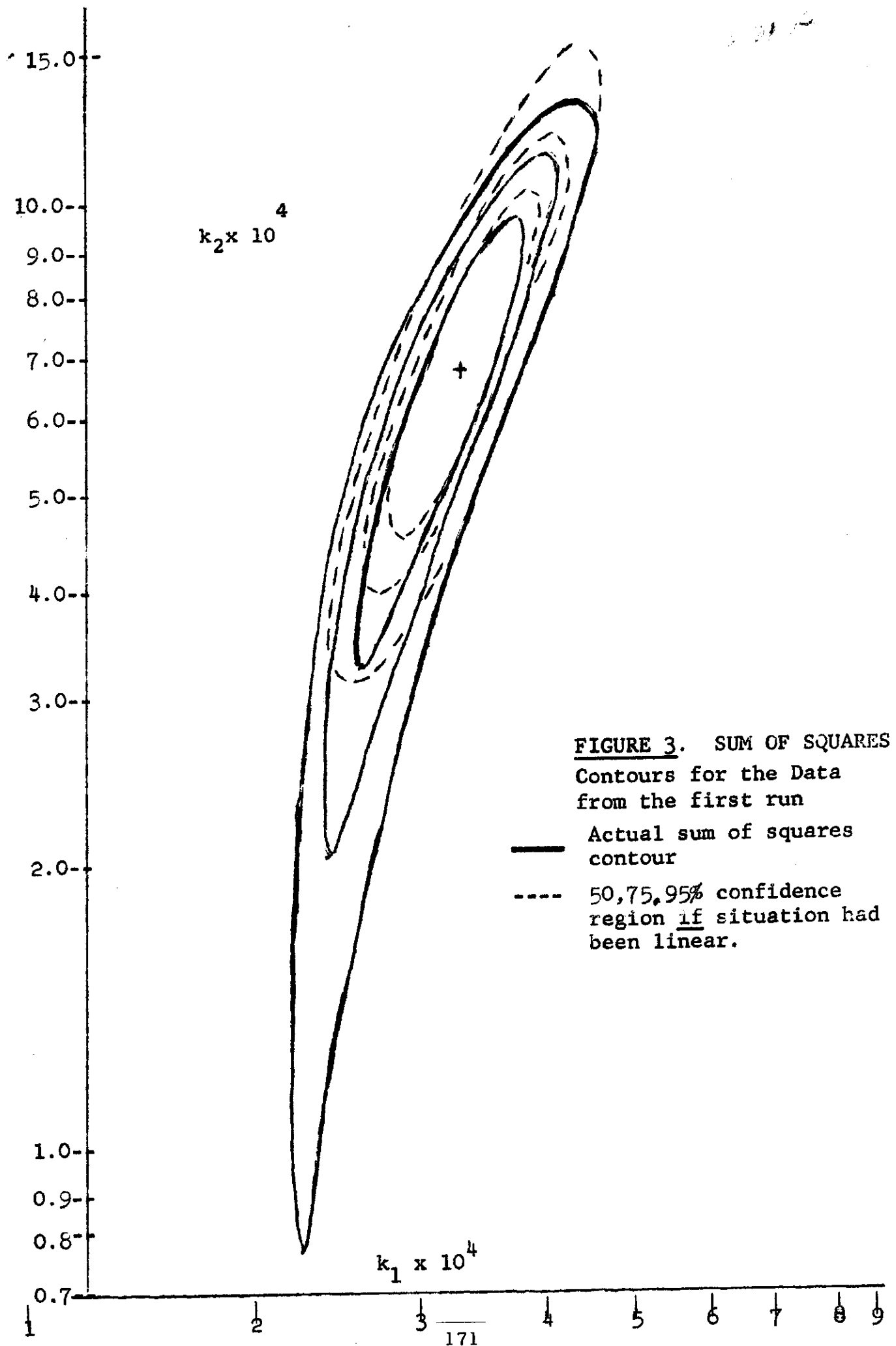
- (1) Does the assumed form of the model fit adequately?
- (2) Assuming it does, are we likely to be misled by the linearization method we are adopting in solving the least squares equations?

Let us consider Question 2 first. For any given run, the estimation situation concerning k_1 and k_2 is completely revealed by a study of the sum of squares surface. On the assumption that the experimental errors are normally distributed with the same variance, the contours of the sum of squares $S(\theta_1, \theta_2) = \sum \{y_i - \eta_i(\theta_1, \theta_2)\}^2$ in the θ_1, θ_2 space are simply the contours of constant likelihood. In these expressions y_i is the observed yield for the time t_i and $\eta(\theta_1, \theta_2)$ is the calculated value for the yield with k_1 and k_2 the values of the constants. The sum of

squares contours for the data from the first run are shown in Figure 3. The contours of the sum of squares surface which would have been appropriate if the situation had been truly linear are shown as dotted lines.

It is seen that the estimation situation does not contain serious abnormalities. In particular the sum of squares surface is reasonably well-conditioned. That is to say it has a clearly defined unique minimum and the contours in the region of this minimum are not seriously attenuated in oblique ridges. The latter situation would indicate serious dependence between the estimates such as might make separate estimation of the constants difficult.

As a general rule, a check of this kind is essential to make sure that the routine calculation not concealing difficulties which may later cause trouble. In the case where there are several parameters, such a study is particularly important. In this case, a complete plot of the likelihood function may be out of the question; however, a survey can be made, by using the method described in [4]. Here the confidence region is "outlined" by determining the points which are the principal axes of the linear theory confidence ellipsoid in a conventional scaling. The theoretical sum of squares and the actual sum of squares are then compared.



The objects are:

- (a) to provide confirmation that a minimum has indeed been found,
- (b) to check linearity (this would be of especial importance in cases where, for example, estimates of the variances and covariances are used in approximate linear theory tests), and
- (c) to show the state of conditioning of the likelihood surface. When there are a number of estimates, it may reveal that certain combinations of the parameters are virtually confounded and therefore it is not really possible to obtain separate estimates.

In Table 4 the residuals from the fitted curves are presented. Relying on the local linearity of the fitted function, the sum of squares of the five residuals from each curve can be used to calculate an estimate of variance based on three degrees of freedom. Pooling the estimates from all the curves, we obtain an estimate of the "within runs" variance of $s^2 = 2.79 \times 10^{-4}$ based on 48 degrees of freedom.

Approximate standard errors for the estimates θ_1 and θ_2 can be obtained in the usual way from the inverse of the sum of squares matrix coming from the final iteration. In this example, we obtain for the sum of squares and cross products

$$C = 10^{-5} \begin{bmatrix} \frac{81.27}{(.05)^2} & \frac{-12.65}{(.05)(.02)} \\ \frac{-12.65}{(.05)(.02)} & \frac{2.52}{(.02)^2} \end{bmatrix}$$

whence

$$C^{-1} = \begin{bmatrix} 14.07 & 28.25 \\ 28.25 & 72.59 \end{bmatrix}.$$

Hence

$$\sigma(\theta_1) = \sqrt{14.07} \sigma = 3.75 \sigma$$

$$\sigma(\theta_2) = \sqrt{72.59} \sigma = 8.52 \sigma$$

where σ is the within run error.

It will be noticed as would be expected that the estimates are somewhat correlated, the correlation coefficient being

$$r = \frac{28.25}{\sqrt{(14.07)(72.59)}} = 0.88$$

These estimates of the standard deviations and the correlation coefficient take account of course only of the within run error and not of errors persisting throughout a run. The latter errors which may be of even greater magnitude than the within run errors arise, for example, from slight misadjustments of the apparatus, discrepancies in the concentrations of solutions, deviations from in-

tended temperature settings, etc. In our analysis in the text we take these errors into account. By employing an estimate of error from the higher order interactions we take account of both the errors arising within runs and between runs.

REFERENCES

1. Box, G. E. P., and Coutie, G. A. (1956), "Application digital computers in the exploration of functional relationships", Proceedings of the Institution of Electrical Engineers, 103, Part B, Supplement No. 1, 100-107.
2. Box, G. E. P. (1957), "Use of statistical methods in the elucidation of basic mechanisms", Bull. Inst. International Statistics, 36, 215.
3. Box, G. E. P. (1960), "The effect of errors in the factor levels and experimental design", Presented at the Inst. International Statistics Meeting, Tokyo, May 30 to June 9, 1960.
4. Box, G. E. P. (1960), "Fitting empirical data", N.Y. Academy of Sciences, 86, 792.
5. Box, G. E. P., and Hunter, J. S. (1961), "The 2^{k-p} fractional factorial designs", Technometrics, 3.
6. Daniel, Cuthbert (1959), "Use of half-normal plots in interpreting factorial two level experiments", Technometrics, 1, 311.
7. Hartley, H. O. (1961), "The modified Gauss-Newton method for the fitting of non-linear regression functions by least squares", Technometrics, 3, 269.