BIOMETRICS *31*, 303–318
June 1975

A FAMILY OF MODELS INVOLVING INTERSECTING STRAIGHT LINES AND CONCOMITANT EXPERIMENTAL DESIGNS USEFUL IN EVALUATING RESPONSE TO FERTILIZER NUTRIENTS[1]

R. L. ANDERSON

University of Kentucky, Lexington, Kentucky 40506, U.S.A.

LARRY A. NELSON

North Carolina State University, Raleigh, North Carolina 27607, U.S.A.

SUMMARY

For many cropping situations, especially in developing countries, quadratic surfaces do not fit the responses of certain crops to fertilizer. Use of the second order designs with standard statistical and economic interpretive techniques may result in costly biases in the estimates of the optimal fertilizer rate. Also, there is a potential pollution problem.

A family of linear-plateau models, consisting of intersecting straight lines, is proposed for fitting fertilizer response data which exhibit a plateau effect. The regression coefficients are easily computed using a desk calculator or computer, and the economic interpretations are simple. Techniques for fitting, parameter estimation, and economic interpretation are described.

For multi-nutrient experiments, a complete factorial experiment with a number of levels of each nutrient is considered to be the best design for both evaluating the model and then estimating the optimal nutrient levels. Preliminary information may provide a basis for deciding which fertilizer nutrients are apt to produce response. In many soil-crop situations, only *NP* or *N* experiments are required, because the other nutrients are already at adequate levels; hence, the amount of experimental material may be redistributed by having fewer factors, but more levels of each factor studied.

Two currently used fertilizer response designs, based on preliminary information on optimal nutrient levels, are described; a one-factor-at-a-time design has the disadvantage of providing no estimate of interaction. Several other designs are suggested. We recommend concentrating several treatment levels in the vicinity of the anticipated optimum. Since the sloping phase of the response pattern is more important than the plateau phase, it should receive more attention when distributing treatment levels.

1. INTRODUCTION

As indicated in a previous paper (Anderson and Nelson [1971]), quadratic models often are used to approximate the response patterns of crops to fertilizer. These models are easily fitted and are well adjusted to the fractional factorial designs used in response surface exploration. Estimates of optimal fertilizer rates are readily obtained by standard techniques (Brown [1956]) and, in addition, exact confidence limits can be computed for the optimal combination of fertilizer nutrient rates.

We have observed that for many crop-soil conditions, particularly in developing nations, the underlying response model is not well-defined; hence, the experimental design must provide for some information to evaluate the response model itself. In many cases, some environmental or management factor may impose a ceiling on the yield; in this case, the

[1] This paper was presented at the 8th International Biometric Conference, Constanta, Romania, August 28, 1974.

top of the true response surface often has the appearance of a plateau. The type of crop, as well as the response measured, has a bearing upon the response pattern, including where the plateau starts if one is present. Response of vegetative growth of some crops to nitrogen (N) may show a plateau starting at a larger value of N than that for yield. Boyd [1970], who conducted studies involving many rates of fertilizer for the express purpose of characterizing the form of the single nutrient response curves for a number of crops, reported that plateaus were reached at low or intermediate rates of fertilizer for a number of different crop-nutrient combinations. Yields of some small grains and potatoes even decreased with large applications of nitrogen; in the case of small grains, lodging may have caused the decline. This response pattern for small grains has also been observed for rice at the International Rice Research Institute, Los Baños, the Philippines.[2]

In the 1971 paper, estimates of residual nutrient levels (δ) were compared when quadratic, square root and exponential response models were used for experimental situations which had yield plateaus. It appeared that the estimates of δ were biased upwards (often by substantial amounts) when the quadratic and square root models were used. Upward biases were also noted in estimated optimal amounts of nutrient to be added; a quadratic-plateau model was considered as a possible replacement. The current paper discusses some new models which have been developed to estimate the optimal rates of fertilizer under experimental space restrictions. The problem discussed in this paper is important because of the spiralling price of fertilizer. Procedures which tend to produce upward biases in estimated optimal fertilizer rates cannot be recommended, particularly for farmers in the developing nations who have great difficulty in obtaining capital to purchase fertilizers. Also, over-fertilization, especially of nitrogen, may have a deleterious effect upon the environment.

2. USE OF QUADRATIC AND SQUARE ROOT MODELS FOR SINGLE NUTRIENT EXPERIMENTS

The bias problem is illustrated by three sets of data for single nutrient experiments. The first set consists of two nitrogen-sugarcane randomized block experiments (three blocks) in Thailand[3], which produced the following mean yields (kg/rai):

	N(kg/rai)			
Place	0	12	24	36
Petburi	5,453	8,028	7,459	7,274
Supanburi	10,654	12,414	11,566	11,887

For each experiment, P and K were applied at respective rates of 12 and 24 kg/rai. A quadratic prediction function, $\hat{Y} = b_0 + b_1N - b_{11}N^2$ was fit by least squares to each set, producing the following values of the b's.

	b_0	b_1	b_{11}	N_0	N_e	$S_{\bar{y}}$
Petburi	5,630	213.2	4.79	22.3	18.8	340
Supanburi	10,843	113.7	2.50	22.7	16.1	289

[2] Personal communication, Dr. Kwanchai A. Gomez, Statistician, International Rice Research Institute, Los Baños.

[3] Data provided by Miss Sanga Duangratana, Statistician, Planning Division, Agriculture Department, Ministry of Agriculture, Bangkok, Thailand.

The maximum value of N is $N_0 = b_1/2b_{11}$. If the ratio of the cost of N to the price of sugarcane is r, then the economic optimal N is $N_e = (b_1 - r)/2b_{11}$. For the above experiments, $r = 0.440/0.0132 = 33.33$. Values of N_e are also presented above. For each experiment there is an apparent upward bias in both N_0 and N_e, because the yields actually decline for $N > 12$. The standard error of each mean yield is presented as $S_{\bar{y}}$.

We consider both of the above to be plateau patterns. A partial explanation for the quadratic bias in plateau response situations was reported in our 1971 paper. If a plateau is reached at low or intermediate levels of added nutrient, b_1 and b_{11} will be biased downward, but the percentage bias of b_{11} is greater than of b_1, hence N_0 and N_e will be biased upwards. The biases in regression coefficients are pronounced when the nutrient being studied is deficient initially and, therefore, the response to the first increment of fertilizer is marked. The parabolic curve is not flexible enough to accomodate the sharp initial rise and subsequent flattening. The top of the arch is to the right of the point where the response to fertilizer ceases. It is also higher than any of the observed yields unless there is decreasing yield at the high rates.

The other two sets of data (Tables 1 and 2) exhibit the plateau effect of nitrogen-corn experiments in the United States. The North Carolina experiments were conducted on Norfolk soil, using an 18-treatment central composite response surface design for three factors (N, P, K); the 22 experiments considered in Table 1 had drought-free soils and

TABLE 1

North Carolina average corn yields and estimated economic optimal values of N using quadratic (Q), square root (S) and linear-plateau (P) models[1]

Location

N	553	554	555	556	559	650	651	652	653	654	655
0	34	12	25	46	24	47	42	41	31	23	55
62.5	78	40	66	79	67	90	71	78	67	62	91
125.0	93	40	89	84	73	111	84	104	77	78	104
187.5	93	35	95	89	76	110	86	106	78	85	112
250.0	95	40	87	85	70	111	84	107	73	82	112
R_0^2(Q)	95.1	64.4	99.9	92.4	91.2	96.8	98.5	98.4	95.5	98.8	98.2
R_0^2(S)	99.1	91.2	94.4	98.7	99.5	96.8	97.7	94.8	98.9	98.0	99.3
R_0^2(P_4)	99.8	95.9	98.7	97.5	98.6	99.9	99.5	99.8	98.8	(98.2)	(96.5)
N_e(Q)	151	96	155	131	133	154	138	161	133	155	160
N_e(S)	114	54	131	81	84	124	93	154	83	126	138
N_e(P_4)	85	63	99	76	72	93	91	108	79	(93)	(96)

N	656	657	658	659	660	661	662	751	752	753	754
0	57	29	64	32	41	29	42	52	28	20	53
62.5	103	72	86	84	79	75	77	95	70	67	86
125.0	115	97	92	108	101	100	97	131	82	96	94
187.5	118	101	97	115	106	101	105	135	87	103	92
250.0	115	102	93	116	103	105	101	139	82	96	91
R_0^2(Q)	95.2	99.0	97.5	98.6	99.7	97.7	99.9	98.5	95.5	99.9	91.2
R_0^2(S)	99.5	96.9	98.1	98.6	95.8	97.3	95.4	95.0	99.2	94.5	99.3
R_0^2(P_4)	99.5	99.3	96.7	(98.4)	99.2	99.5	(98.0)	99.2	98.9	98.9	99.3
N_e(Q)	147	165	117	172	158	165	159	177	147	163	125
N_e(S)	105	165	64	184	142	160	147	243	106	162	73
N_e(P_4)	81	103	84	(98)	104	99	(106)	119	82	104	74

[1] Yields rounded to nearest bu/acre and N in lbs./acre; computations based on non-rounded data. R_0^2=percent variation explained by regression. N_e is economic optimal N based on r=0.12/lb N. P_4 refers to Model IV_1 (except III_1 for 554); parentheses refer to four experiments for which Model VII_{12} is better. The best fits for the latter experiments were:

Experiment	654	655	659	662
R_0^2	99.8	99.9	99.9	99.7
N_e	148	162	146	143

Standard errors and no. of observations (in parentheses) for successive N-means: 3.40(8); 2.41(16); 2.15(20); 2.41(16); 2.78(12).

TABLE 2

TENNESSEE AVERAGE CORN YIELDS AND ESTIMATED OPTIMAL VALUES OF N USING QUADRATIC (Q), SQUARE ROOT (S) AND LINEAR-PLATEAU (P) MODELS[1]

	Knoxville					Jackson				
N	1962	1963	1964	1965	1966	1962	1963	1964	1965	1966
0	44.6	45.1	60.9	59.4	63.0	46.5	29.3	28.8	25.8	23.4
67	73.0	73.2	75.9	67.4	61.3	59.0	55.2	37.6	47.6	45.2
134	75.2	89.3	83.7	69.6	63.4	71.9	77.3	55.2	60.5	53.3
201	83.3	91.2	84.3	62.9	59.4	73.1	88.0	66.8	70.2	61.3
268	78.4	91.4	81.8	67.0	57.8	74.5	89.4	67.0	68.0	57.6
335	80.9	88.0	84.5	61.2	58.3	75.5	87.0	67.8	73.0	59.8
R^2(Q)	0.895	0.977	0.922	0.561	--	0.977	0.998	0.969	0.982	0.966
R^2(S)	0.976	0.972	0.968	0.710	--	0.955	0.954	--	0.977	0.977
R^2(P)	0.986	0.995	0.989	0.403	--	0.996	0.998	0.991	0.991	0.993
P-Model	VI_{13}	IV_1	IV_1	III_1	I	IV_2	IV_2	III_3	VI_{13}	VI_{13}
N_e(Q)	218	219	209	118	0	238	250	284	260	230
N_e(S)	189	217	170	60	0	349	568	--	595	261
N_e(P)	201	107	101	67	0	147	162	201	201	201

[1] Yields in q/ha and N in kg/ha. Economic optimal rates (N_e) based on r = 0.0244 per kg (1.635 for 67 kg). Square root model could not be used for J-64. Yields are averages of nine observations; approximate standard error of each average is 1.7. R^2 = proportion of variation explained by regression.

responded only to N. Descriptions of sites, experimental procedures and yield data are reported in our 1971 paper and by Baird and Fitts [1957]. Estimates of N_e were obtained for three models: quadratic (described above); square root (N is replaced by $\sqrt{N}$ in the quadratic); linear-plateau (to be described later). The proportion of the treatment sum of squares attributed to regression (R^2) was calculated for each experiment and model. For many experiments, the values of N_e for the quadratic and square root models appear high, especially when one compares N_e with an estimate based on a careful comparison of the mean yields and their standard errors.

The Tennessee data (Table 2) are taken from experiments of Engelstad and Parks [1971], conducted for five years at two locations (near Knoxville and Jackson). For the Knoxville data, the quadratic model seemed to overestimate N_e for 1963–65 and the square root model in 1963–64. For the Jackson data, the square root model could not be used in 1964 and vastly overestimated N_e in 1962, 1963 and 1965; the quadratic model also seemed to overestimate N_e for 1962–1965.

3. DEVELOPMENT OF THE LINEAR-PLATEAU FAMILY OF SINGLE NUTRIENT MODELS

After it became apparent that the calculated optimal rates obtained by fitting the quadratic and square root models to these three sets of data generally were biased upward, a new model which would mitigate these biases was sought. The analysis of such a model might have these requirements:

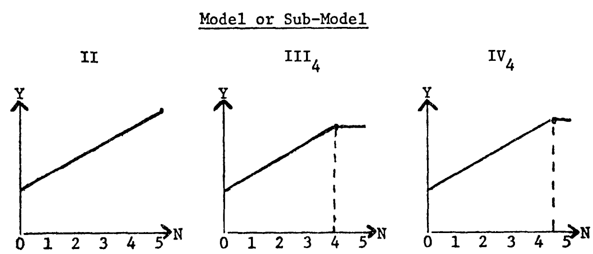
(1) lead to reasonably accurate estimates of the optimal fertilizer rates for various decision rules;
(2) produce a satisfactory goodness-of-fit to the data;
(3) be easily adapted to obtain results based on the average of a number of experiments;

(4) be amenable to easy calculation.

Condition (3) almost dictates the use of a model which is linear in the parameters. Condition (4) is included because the resulting model would be used in developing nations where computing equipment is often not very sophisticated if it exists at all. The spline models (Fuller [1969] and Wold [1974]) were rejected because they require considerable calculation, and our data did not have enough levels in the region of the join points.

It was deemed impossible to use a single continuous model which would apply to and fit all response situations and produce unbiased estimates of optimal rates. Instead we have developed a system of models belonging to a "linear-plateau family", which generally involve intersecting lines rather than a continuous curve. The fertilizer response data are used to estimate the parameters of each of the models of the family, to test the significance of these estimates and to decide which model of the family provides the best fit. The name *linear-plateau* implies a region of linear response (with possibly more than one slope) and a plateau. The linear-plateau models have some features similar to that developed for a graphical fitting approach by Waugh *et al.* [1973].

Diagrams are presented in Figures 1 and 2 of linear-plateau patterns for experiments using six equally spaced nutrient levels. A brief description of each diagram is included below the diagrams. In these diagrams, the horizontal axes represent coded nutrient levels ($N = 0$ for check; $N = 1, 2, 3, 4, 5$ for the equally spaced increments of nutrient), and the



II: Single sloping line.

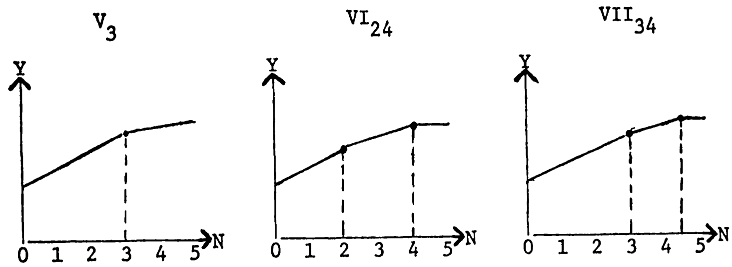
III_4: Single sloping line intersecting a plateau at $N = 4(b_1>0)$. Similar sub-models are III_3 (intersection at N=3), III_2 (intersection at N=2) and III_1 (intersection at N=1).

IV_4: Single sloping line intersecting a plateau between N=4 and $5(b_1>b_3>0)$ or two sloping lines; insufficiency of data prevents distinguishing between the two, but we choose to use the sloping line-plateau. Similar sub-models are IV_3 (intersection between N=3 and 4), IV_2 (intersection between N=2 and 3) and IV_1 (intersection between N=1 and 2).

FIGURE 1

DIAGRAMS AND DESCRIPTIONS FOR MODELS II–IV

Model or Sub-Model



V_3: Two sloping lines intersecting at $N=3(b_1>b_2>0)$. Two other sub-models are V_2 (intersection at $N=2$) and V_1 (intersection at $N=1$).

VI_{24}: Two sloping lines intersecting at $N=2$ with the second sloping line intersecting a plateau at $N=4(b_1>b_2>0)$. Two other sub-models are VI_{14} (intersections at $N=1$ and 4) and VI_{13} (intersections at $N=1$ and 3). If the second sloping line were to connect two adjacent points, one could use either IV or VI; we prefer IV, especially if the cost/price data suggest that addition in the second phase of VI would be unprofitable.

VII_{34}: Two sloping lines intersecting at $N=3$ and the second sloping line intersecting a plateau between $N=4$ and $5(b_1>b_2>b_3>0)$. Similar sub-models are VII_{24} (intersections at $N=2$ and between $N=4$ and 5), VII_{14} (intersections at $N=1$ and between $N=4$ and 5), VII_{23} (intersections at $N=2$ and between $N=3$ and 4), VII_{13} (intersections at $N=1$ and between $N=3$ and 4) and VII_{12} (intersections at $N=1$ and between $N=2$ and 3).

FIGURE 2
DIAGRAMS AND DESCRIPTIONS FOR MODELS V–VII

vertical axes represent average yield (Y). The mathematical models for these diagrams are of the type

$$Y = \beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e,$$

where Y is the average yield; β_0 is the intercept or check yield (X_0 is a vector of ones); β_1 and β_2 are the slopes of the sloping lines; β_3 is the distance of the plateau above the value of Y at the last design point before the start of the plateau; e is the residual or experimental error. Note that these models have a maximum of four regression coefficients. The prediction equation is

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 ,$$

where b_i is the experimental estimate of β_i. We have omitted Model I from Figure 1, because it is simply a plateau, $\hat{Y} = \bar{Y}$; this is the case for which there is no response to added nutrient.

Diagrams with one sloping line are presented in Figure 1; those with two sloping lines are in Figure 2. All models except II and V have plateaus. Note that there are no more than two sloping lines plus at most one plateau for any model. Models III–VII have a number of sub-models. For these sub-models, a single subscript on the model number refers to the location of the sloping line-plateau intersection (III_i or IV_i) or the intersection of two sloping lines (V_i). For Models VI and VII, two subscripts (ij) are needed; the first subscript (i) refers to the location of the intersection of the two sloping lines, and the second subscript (j) refers to the location of the intersection of the second sloping line and a plateau. The intersection of two sloping lines will always be at a design point. If the right-hand sloping line intersects a plateau, the intersection may be at a design point or between design points; the data are used to determine which is the case, i.e. test if b_3 is significantly greater than zero. In general, it is required that $b_1 > b_2 > b_3 > 0$; if b_2 or b_3 is negative, either a quadratic response model should be used or the data for declining yields should be discarded.

If b_3 is not significantly greater than zero for Model IV_i, Model III_i should then be considered. If b_3 is not significantly greater than zero for Model VII_{ij}, there are two possibilities:

(i) if $j > i + 1$, consider Model VI_{ij};
(ii) if $j = i + 1$, consider Model IV_i.

Many tests for choice of model are more complicated than a mere test of, for example, $\beta_3 = 0$. For example, one might wish to test if Model VII_{12} is superior to Model VI_{13}. Since a major objective in any model building is to use as few parameters as possible, one would favor VI over VII unless the latter is definitely superior. The appropriate test of VII_{12} versus VI_{13} for the six-level experiment is to test $\beta_3 = \beta_2$ for VII_{12}. This test would be based on the contrast $l = -Y_1 + 2Y_2 - \bar{Y}_3$, where Y_i is the average yield for $N = i$ and $\bar{Y}_3$ is the average yield for $N = 3$, 4 and 5. Note that if in VII_{12} $\beta_3 = \beta_2$, then Model VII_{12} is the same as Model VI_{13}. Similar tests would compare VII_{13} with VI_{14}, VII_{23} with VI_{24}, VII_{14} with V_1, VII_{24} with V_2 and VII_{34} with V_3.

It is recognized that one might obtain a slightly better fit for the models with two sloping lines (V, VI and VII in Figure 2) by allowing the sloping lines to intersect between design points. However, experience has indicated that this slight improvement is not worth the added programming complications.

In order to indicate the simplicity of the computations needed to test the models and estimate the β's, the Y, X_1, X_2 and X_3 vectors are presented in Table 3 for each model and sub-model mentioned in Figures 1 and 2.

The statistical analysis for all of these models and sub-models can be performed readily without the use of sophisticated computing equipment. Sets of least squares equations can be developed for calculating the b's and their standard errors for experiments involving various numbers of levels and nutrients. The yields obtained in a given experiment would be substituted into these equations and the parameter estimates calculated efficiently using only a pocket or desk calculator. This extends the usefulness of the method to developing countries and isolated experiment stations. If an electronic computer is available, it may be programmed to fit the entire series for each number of levels of nutrient and to print out regression statistics and information on goodness-of-fit. It is assumed that enough

TABLE 3

YIELD AND X-VECTORS FOR MODELS AND SUB-MODELS OF FIGURES 1 AND 2

	Model II	Model III				Model IV							
		Sub-model				Sub-model							
		4	3	2	1	4		3		2		1	
Y	X_1	X_1	X_1	X_1	X_1	X_1	X_3	X_1	X_3	X_1	X_3	X_1	X_3
Y_0	0	0	0	0	0	0	0	0	0	0	0	0	0
Y_1	1	1	1	1	1	1	0	1	0	1	0	1	0
Y_2	2	2	2	2	1	2	0	2	0	2	0	1	1
Y_3	3	3	3	2	1	3	0	3	0	2	1	1	1
Y_4	4	4	3	2	1	4	0	3	1	2	1	1	1
Y_5	5	4	3	2	1	4	1	3	1	2	1	1	1

	Model V						Model VI					
	Sub-model						Sub-model					
	3		2		1		24		14		13	
Y	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2
Y_0	0	0	0	0	0	0	0	0	0	0	0	0
Y_1	1	0	1	0	1	0	1	0	1	0	1	0
Y_2	2	0	2	0	1	1	2	0	1	1	1	1
Y_3	3	0	2	1	1	2	2	1	1	2	1	2
Y_4	3	1	2	2	1	3	2	2	1	3	1	2
Y_5	3	2	2	3	1	4	2	2	1	3	1	2

Model VII

	Sub-model																	
	34			24			14			23			13			12		
Y	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
Y_0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Y_1	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
Y_2	2	0	0	2	0	0	1	1	0	2	0	0	1	1	0	1	1	0
Y_3	3	0	0	2	1	0	1	2	0	2	1	0	1	2	0	1	1	1
Y_4	3	1	0	2	2	0	1	3	0	2	1	1	1	2	1	1	1	1
Y_5	3	1	1	2	2	1	1	3	1	2	1	1	1	2	1	1	1	1

design points are replicated to provide a good estimate of the error variance. A bulletin is being considered to present the detailed computational procedures for experiments having up to seven factor levels.

The economic implications of the linear-plateau family are quite simple and straightforward. In the rules described below, r refers to the cost of a unit of fertilizer nutrient divided by the price of a unit of crop produced. The value of r must be coded in the same unit as the b's. For example, for the Tennessee data, the coded N rates represent 67 kg units; therefore, r must be multiplied by 67 when comparing it with the coded b's. The rules for the addition of coded amounts (N) of nutrient are as follows:

I: (and other models with $b_1 < r$): add none or only maintenance rate.

II: If $b_1 > r$, add largest N in experiment (further experiments needed to establish exact optimum).

III_j : If $b_1 > r$, add $N = j$ ($N = 4$ for III_4 , etc.).

IV_j : If $b_1 > r$ and $b_3 > 0$, add $N = j + b_3/b_1$.

V_i : (1). If $b_1 > r > b_2$, add $N = i$ ($N = 3$ for V_3).

(2). If $b_2 > r$, same as II.

VI_{ij} : (1). If $b_1 > r > b_2$, add $N = i$ ($N = 2$ for VI_{24}).

(2). If $b_2 > r$, add $N = j$ ($N = 4$ for VI_{24}).

VII_{ij} : (1). If $b_1 > r > b_2$, add $N = i$.

(2). If $b_2 > r$ and $b_3 > 0$, add $N = j + b_3/b_2$.

Models V, VI and VII also permit the option of fertilizing only to the intersection of the sloping lines, even though it may be profitable to fertilize in the region of the second sloping line (up to the intersection with the plateau for Models VI and VII). This fertilization strategy could be useful in developing countries where it may be deemed desirable to fertilize only in the region of maximum percentage return to capital investment. The family of linear-plateau models lends itself readily to economic programming because of the existence of these decision points and the relatively simple patterns of fertilization strategies throughout the entire treatment range.

To illustrate the computing procedures, we will consider the J-65 data of Table 2. Since the computer manual mentioned above would contain procedures for each of the models and sub-models, only the procedures for the two best-fitting models (VII_{12} and VI_{13}) will be included here. It is assumed that a preliminary analysis of the experimental data has produced an estimate of the standard deviation; in this case $s = 1.7$. Model VII_{12} will be designated as VII and Model VI_{13} as VI.

(*1*) *Compute the vector products:* $g_i = X_i'Y$.
  (i) For both models, $g_0 = \sum Y = 345.1$ and $g_1 = g_0 - Y_0 = 319.3$.
  (ii) For VII, $g_2 = g_1 - Y_1 = 271.7$ and $g_3 = 3\bar{Y}_3 = 211.2$.
  (iii) For VI, $g_2 = Y_2 + 6\bar{Y}_3 = 482.9$.

(*2*) *Compute the b's:*
  (i) For both models, $b_0 = Y_0 = 25.8$.
  (ii) For VII, $b_1 = Y_1 - Y_0 = 21.8$; $b_2 = Y_2 - Y_1 = 12.9$; $b_3 = \bar{Y}_3 - Y_2 = 9.9$.
  (iii) For VI, $b_1 = (-16g_0 + 29g_1 - 7g_2)/16 = (-16Y_0 + 13Y_1 + 6Y_2 - 3\bar{Y}_3)/16 = 22.3625$; $b_2 = (-7g_1 + 5g_2)/16 = (-7Y_1 - 2Y_2 + 9\bar{Y}_3)/16 = 11.2125$.

(*3*) *Compute standard errors of the b's:*
  (i) For VII, $s_1 = s_2 = 2.4$; $s_3 = s\sqrt{4/3} = 2.0$, where $s_i = s.e.(b_i)$.
  (ii) For VI, $s_1 = s\sqrt{1.8125} = 2.3$; $s_2 = s\sqrt{0.3125} = 0.95$.

(*4*) *Compute the residual sum of squares (SSE) and mean square (MSE):*
  (i) $\sum Y^2 = 21472.69$; $\text{SSE} = \sum Y^2 - \sum b_i g_i$; $\text{MSE} = \text{SSE}/f$, where $f = 2$ for VII and $f = 3$ for VI.
  (ii) For VII, MSE = 6.29.
  (iii) For VI, MSE = 4.75.

(*5*) *Compute* R^2: $R^2 = 1 - [\text{SSE}/\sum (Y - \bar{Y})^2]$.
R^2 (VII) = 0.9923 and R^2 (VI) = 0.9912.

(*6*) *Compare Models:*
Since MSE (VII) is larger than MSE (VI), it seems reasonable to use Model VI_{13}. The test of $\beta_3 = \beta_2$ for VII_{12} produces

$$t = \frac{2Y_2 - Y_1 - \bar{Y}_3}{s\sqrt{16/3}} = \frac{3}{3.9} = 0.8;$$

this also leads to use of Model VI_{13}.

(*7*) The notation is complicated because N is coded in units 0, 1, 2, $\cdots$ for the above computations and in Figures 1 and 2 and the rules for determining the optimum; however, the experimenter desires the optimum to be quoted in terms of amount of actual nutrient. We will use here the notation N_e for the optimal amount of actual nutrient to be added. Assuming Model VI_{13} is used, $b_1 > b_2 > r = 1.635$; hence, we add coded $N = j = 3$ and $N_e = 67(3) = 201$. If Model VII_{12} had been used, $b_1 > b_2 > r$ and $b_2 > b_3$; hence one could add coded $N = j + b_3/b_2 = 2.767$ and $N_e = 185$.

4. USE OF LINEAR-PLATEAU MODELS FOR THE DATA IN TABLES 1 AND 2

As indicated in Section 2, the North Carolina and Tennessee corn yield data (Tables 1 and 2) exhibit definite plateau patterns. For the North Carolina data (Table 1), Model IV_1 gave the best fit of the linear-plateau models for 17 locations, Model III_1 for 1 location and Model VII_{12} for 4 locations, using a cost/price ratio of 0.12. These results for single locations in a given year are not too useful for general fertilizer recommendations. For this purpose, it was decided to determine N_e for the average yields based on these 22 locations. These average yields are:

N	0	1	2	3	4
Y	37.52	76.51	93.03	96.59	94.88
n	176	352	440	352	264

The values of n are the number of observations for each average.

Since the increase in yield from $N = 2$ to $N = 3$ is less than the cost/price ratio for 62.5 lbs/acre (7.5), we have decided to use Model IV_1 instead of Model VII_{12} ; also the former was the predominant model for individual locations. In this case

$$b_0 = Y_0 = 37.52; \qquad b_1 = Y_1 - Y_0 = 38.99; \qquad b_3 = \bar{Y}_2 - Y_1 = 18.17.$$

The standard errors of b_1 and b_3 are based on the location-to-location variability of the two sets of differences; $s(b_1) = 1.51$ and $s(b_3) = 2.13$. The optimal N would be

$$N_e = 62.5(1 + b_3/b_1) = 91.6 \text{ lbs./acre.}$$

The above N_e is approximately the same as the average N_e , averaged over all locations if Model IV_1 were used; this average is 91.3 ± 3.0, where the standard error is computed from the location-to-location variation in the N_e's. The two estimates will not be identical because Ave $(b_3/b_1) \neq$ Ave (b_3)/Ave (b_1). If we use the Model VII_{12} results for locations 654, 655, 659, and 662, the average N_e is 100.7 ± 5.8. If the quadratic and square root models had been used on the location averages, the optimal N would have been: $N_e(Q) = 151$ and $N_e(S) = 118$ for $r = 0.12$. Averaging over locations, $\bar{N}_e(Q)$ is 148.27 ± 4.14 and $\bar{N}_e(S)$ is 124.23 ± 9.55. As indicated above, it appears that N_e should be no more than 125 lbs./acre, and it is doubtful if this much would be profitable; therefore, the quadratic estimate of N_e seems to be unacceptable.

In summary, it appears that N_e lies somewhere between 62.5 and 125 lbs./acre. The same uncertainty exists for each of the three years. These results suggest that the experimental program, after the first year, should have included some nitrogen applications between 62.5 and 125 lbs./acre in order to more accurately determine where the plateau starts.

For the Tennessee data (Table 2), the average yields are:

N	0	67	134	201	268	335
Knoxville	54.60	70.16	76.24	76.22	75.28	74.58
Jackson	30.76	48.92	63.64	71.88	71.30	72.62

It is not possible to pool these two sets of data, because the two sites seem to have different fertilizer requirements. Standard errors are very imprecise because there are only five observations for each N at each site. The Knoxville data are difficult to analyze because Model I is appropriate for 1966. If we use $b_1 = b_3 = 0$ for 1966 and Model III_1 or IV_1 for the other years, the average b_i, N_e and their standard errors are

$$\bar{b}_1 = 15.9 \pm 5.6; \qquad \bar{b}_3 = 6.18 \pm 3.09; \qquad \bar{N}_e = 71.4 \pm 19.2.$$

If Model VI_{13} is used for 1962, $\bar{N}_e = 95.2 \pm 32.6$. Similarly, $\bar{N}_e(Q) = 152.8 \pm 42.6$; $\bar{N}_e(S) = 127.2 \pm 41.5$. In this case it appears that the quadratic model definitely overestimates N_e. There is uncertainty as to where the optimum is between 67 and 134 kg/ha; again the experimenter should introduce treatments between these two levels in subsequent experiments.

The conclusions at Jackson can be made somewhat more precisely because there were no Model I data. Using Model IV_2 or III_3,

$$\bar{b}_1 = 16.44 \pm 2.06, \qquad \bar{b}_3 = 7.67 \pm 1.97, \qquad \bar{N}_e = 165.8 \pm 9.5.$$

If the Model VI_{13} results are used for 1965 and 1966,

$$\bar{N}_e = 182.4 \pm 11.6.$$

Similarly, $\bar{N}_e(Q) = 252.4 \pm 9.4$; the square root model could not be used here. Again the quadratic estimate is far too large; the results would have been more precise if additional levels of nutrient between 134 and 201 kg/ha had been used.

5. FITTING QUADRATIC, SQUARE ROOT AND LINEAR-PLATEAU MODELS TO MULTI-NUTRIENT DATA

A fourth set of data (Table 4) illustrates a plateau pattern for an NPK experiment on corn conducted at the Praputhabat Agricultural Experiment Station, Saraburi, Thailand, in 1962.[4] There was significant response to K only at the 0 level of N. A three-variable quadratic surface was modified because the coefficient of K^2 was positive with an estimated standard error twice the estimated coefficient; the modified quadratic prediction model was:

$$\hat{Y}(Q) = 279.17 + 166.018\,N - 35.706\,N^2 + 85.954\,P - 27.141\,P^2$$
$$+ 18.319\,NP + 60.474\,K - 26.539\,NK - 6.437\,PK,$$

where N, P and K are in coded units (0, 1, 2, 3). The corresponding square root prediction model (omitting the coefficient of K) was:

$$\hat{Y}(S) = 270.01 - 56.111\,N + 201.161\sqrt{N} - 70.577\,P + 122.064\sqrt{P}$$
$$+ 59.165\sqrt{NP} + 100.785\sqrt{K} - 69.361\sqrt{NK} - 14.928\sqrt{PK}.$$

The prices per kg were as follows:

Corn (\$0.117), N(\$0.440), P(\$0.352), K(\$0.176).

Therefore, the cost/price ratios for the coded units were:

$$r_n = 30.08, \quad r_p = 10.50 \quad \text{and} \quad r_k = 9.99.$$

The resulting optimal fertilizer combinations (in kg/rai) had $K = 0$ and

$$N_e(Q) = 19.81, \quad N_e(S) = 21.24; \quad P_e(Q) = 7.77, \quad P_e(S) = 6.34.$$

[4] Data provided by Dr. Sorasith Vatcharothayan, Head, Department of Soil Science, Kasetsart University, Bangkok, Thailand.

These results for both N and P seem biased upwards because the predicted optimal yields

$$\hat{Y}(Q) = 629.15; \qquad \hat{Y}(S) = 615.05$$

exceed by a considerable amount any mean in Table 4, which indicates little, if any, response beyond 8 kg N/rai and 3.49 kg P/rai. From Table 4, it appears that one can obtain a yield of about 446 kg/rai with $N = 0$, $P = 0$ and $K = 13.28$ kg/rai.

For $N > 0$ and $P > 0$, the optimal K is 0; in this case, it appears that a combination of linear-plateau Model III_1 for P and Model IV_1 for N would be appropriate. Assuming no interaction, the model would be

$$\hat{Y} = b_0 + b_{1P}X_{1P} + b_{1N}X_{1N} + b_{3N}X_{3N} .$$

The following sets of data are pooled to estimate the b's: $K = 0, N = 0$, all P; $P = 0$, all K, $N > 0$; $N = 1$, all $K, P > 0$; $N > 1$, all $K,P > 0$. The estimating equations and estimates are:

b_0	b_{1P}	b_{1N}	b_{3N}	Yields	
1	0	0	0	287.16	$b_0 = 287.2 \pm 28.1$
1	0	1	0	438.59	$b_{1P} = 93.4 \pm 26.4$
1	1	1	0	532.01	$b_{1N} = 151.4 \pm 33.8$
1	1	1	1	580.06	$b_{3N} = 48.0 \pm 22.9$

In this case $P_e = 3.5$ kg/rai and $N_e = 8(1 + 48/151) = 10.5$ kg/rai, with an approximate standard error of 1.34 (considering N_e as a ratio estimate).

The estimated gain would be

$$580(0.117) - 10.5(0.440) - 3.5(0.352) = \$62.01 \text{ per rai}$$

TABLE 4

SELECTED MEAN CORN YIELDS FROM 4 $\times$ 4 $\times$ 3 NPK EXPERIMENT, SARABURI, THAILAND[1]

	P					
NK	0	1	2	3	Ave.	
00	298.24	320.34	254.10	275.95	287.16	
01	309.44	367.49	397.04	293.17	341.78	
02	446.42	517.44	414.62	419.92	449.60	
					Ave. (P>0)	
1-	437.66	541.46	530.94	523.63	532.01	
2-	431.33	579.03	569.30	587.40	578.58	580.06
3-	446.79	577.95	576.56	590.14	581.55	
Ave.(N>0)	438.59	566.15	558.93	567.06	564.05	

Standard errors of means

NPK	NP	P(N>0)	N(P>0)	(N,P>0)
56.2	32.5	18.7	18.7	10.8

[1] Yields in kg/rai; N in units of 8 kg/rai; P in units of 3.49 kg/rai and K in units of 6.64 kg/rai. NPK means based on three replications.

as compared to \$49.89 using only K. The estimated gain would be \$62.16 per rai using the quadratic model and \$60.39 using the square root model; however, these gains are overestimated because the predicted optimum yield ($\hat{Y}$) is too large. If one uses 590 as the maximum yield, the gain using Q would be \$57.58 and using S would be \$57.46. Gains from all models would need to be multiplied by 6.25 to convert to a hectare basis. One might consider penalizing the Q and S models for the harm to the environment of overfertilization.

For experiments in which there are two or more nutrients varied factorially and which exhibit a plateau response pattern, the procedures outlined in Section 3 for a single nutrient can be applied individually to means, the choice of which depends upon the response pattern as reflected in the ensemble of means and corresponding analyses of variance. In the most complicated situation, it would be necessary to make estimates of the optimal rates on individual rows and columns, assuming only two variables, a procedure which is not too efficient. The optimal values would then be connected by contours connecting the row and column optima and the contours then plotted. The point where the two contours intersect is taken for the optimal combination. If there is no interaction so that one can average over levels of the other factor, a more stable estimate should be obtained.

Since a common plateau may extend over several rows and columns, decisions made on individual rows and columns cannot be completely independent. Independent estimates for the various rows and columns may be made first and then some adjusting in individual row models may be necessary to make each row or column conform more to adjacent rows or columns.

Our methods work well in multi-nutrient situations where there are not sizeable interactions in the vicinity of the optimum (such as in Table 4). More complicated response patterns will need additional study. We hope to develop a set of general rules for the multi-nutrient case in a manner similar to that developed for single nutrient models.

6. DESIGN OF EXPERIMENTS FOR CHOICE OF MODEL AND DETERMINATION OF OPTIMAL RATES OF FERTILIZER

Most soil fertility experiments involve the major nutrients, N, P and K. If one were to make an assessment of the overall importance of these nutrients in decreasing order of response occurrence, the ordering would be $N > P > K$. There are many crop-locations in which K may not produce response. Preliminary information should be used to control the size of experiments by not varying factors which are known to be at adequate levels. If the model is not well known, the ideal design arrangement may be a complete factorial experiment with many levels of each factor (say 5–7). If possible, the levels should be spaced so that there is some concentration of treatment points in the region of the expected optimal rate. This is not impractical if the number of nutrient factors is not large (say two). As mentioned above, preliminary information may serve as the basis for exclusion of one or two nutrient factors from the design.

The biases in optimal rate estimates based on quadratic and square root models have led some researchers to believe erroneously that the 0-level treatment is so remote from the optimum that it is not needed in newly designed experiments. We have found cases where all levels in such experiments were set higher than the optimum; therefore, significant response was not obtained to any of the treatments in these new experiments. If a check level had been included in these experiments, Model III_1 or IV_1 might have been appropriate. Until some evaluation of the relevant ranges in response is available, a check treatment should be included in fertilizer experiments.

The North Carolina State University International Soil Fertility Evaluation and Improvement Program in Latin America obtains preliminary information about optimal fertilizer rates on each crop-soil by the following procedures:

(i) Soil tests, accompanied by soil calibration with field results;

(ii) Greenhouse pot trials and exploratory 2^3 factorials of N, P and K.

These preliminary trials indicate if any nutrients can be ignored or added at low maintenance levels. They also provide estimates of the proper levels for each of the important factors in the subsequent randomized complete block experiments which utilize a sequential one-factor-at-a-time approach, with the other factor(s) held at constant (preliminary optimal) levels.

Assuming preliminary information suggests that the respective optimal rates of N, P and K are 24, 12 and 24 kg/ha, an experiment might be conducted using the following 14 treatments:

Treatment	1	2	3	4	5	6	7	8	9	10	11	12	13	14
N	0	12	24	36	48	24	24	24	24	24	24	24	24	0
P	12	12	12	12	12	0	24	36	48	12	12	12	12	0
K	24	24	24	24	24	24	24	24	24	0	12	36	48	0

Treatments 1 through 5 should provide a good picture of the model appropriate for nitrogen response and an estimate of the optimal rate of N. Treatments 6, 3, 7, 8 and 9 provide the phosphorus response information and treatments 10, 11, 3, 12 and 13 the information for potassium. A check treatment (14) probably should be included. This information would be used to center the design for the soil fertility trial conducted the following year. This cycle may be repeated for several years for a given crop-soil until the estimated optimal levels of all factors stabilize.

The arrangement described above does not provide information on interactions but it might be argued that much of the interaction occurs in comparing the results for 0 and 1 applications, which are below the preliminary optima. The data from Saraburi, Thailand (Table 4) demonstrate this type of interaction ($N \times K$) which is not important in the region of the optimum. It should be added that although a complete factorial with many levels of each factor would have been preferred in the above case, space in the farmer-cooperative fields in Latin America is extremely limiting. However, during some phase of the cycle, the design should provide for estimation of interacting effects.

The following exploratory experimental design involving two nutrients is being used in the highlands of Ecuador where space is severely limited.[5] Greenhouse and soil test studies have provided initial estimates of the optimal levels of N and P required for a specific crop-soil. Suppose these estimates of optimal levels of N and P are each coded rates of 3. The design points are as follows:

Treatment	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
N	0	1	2	3	4	5	6	3	3	3	3	3	3	0	1	2	4	5	6	0	6
P	0	1	2	3	4	5	6	0	1	2	4	5	6	3	3	3	3	3	3	6	0

The one-factor-at-a-time approach provides the 13 design points for $P = 3$ and for $N = 3$; treatments along the diagonal (points 1–7) are included to check for interaction and the

[5] We acknowledge the efforts of Dr. Sam Portch of North Carolina State University in the development of this design.

stability of (3, 3) as the optimal point. Treatments 20 and 21 might be included to provide a complete 3×3 factorial (0, 3, 6 levels of N and P), which would prove useful should the model prove to be truly of second order. The experiment probably will be conducted several years in succession, with slight adjustments in the rates until the optimal levels appear to stabilize.

If the experimenter has good reason to believe that there is one very important nutrient (X_1) and one less important (X_2), the experimental design could be a 7×3 factorial. If there are two less important nutrients, (X_2 and X_3), an incomplete 3×3 factorial could be used for the latter such as:

X_2	0	1	2	1	1	0	2
X_3	0	1	2	0	2	1	1.

If Model IV_1 were appropriate for both X_2 and X_3 with no interaction, the model would be

$$\hat{Y} = b_0X_0 + b_{21}X_{21} + b_{23}X_{23} + b_{31}X_{31} + b_{33}X_{33} .$$

The X-vectors would be

	X_0	X_{21}	X_{23}	X_{31}	X_{33}
Y_{00}	1	0	0	0	0
Y_{01}	1	0	0	1	0
Y_{10}	1	1	0	0	0
Y_{11}	1	1	0	1	0
Y_{12}	1	1	0	1	1
Y_{21}	1	1	1	1	0
Y_{22}	1	1	1	1	1

Using all seven points, there would be two degrees of freedom for goodness-of-fit (test of interaction). The variance of b_0 would be $V(b_0) = 11\sigma^2/15$; for others $V = 14\sigma^2/15$. If Y_{22} were omitted, $V(b_{23})$ and $V(b_{33})$ would increase to $7\sigma^2/4$, $V(b_{21})$ and $V(b_{31})$ to σ^2 and $V(b_0)$ to $3\sigma^2/4$. Hence the extra point not only furnishes additional information on goodness-of-fit but considerably reduces the variances of the plateau estimates (b_{23} and b_{33}). Each of these design points might have a complete 5–7 level experiment for the X_1-factor.

UNE FAMILLE DE MODELES A SEGMENTS DE DROITE ET PLANS EXPERIMENTAUX UTILES DANS L'EVALUATION DE LA REPONSE A DES ENGRAIS

RESUME

Dans le cas de nombreuses cultures, en particulier dans les pays développés, des formes quadratiques n'ajustent pas bien les réponses de certaines cultures aux engrais. L'utilisation des plans du second ordre à interprétations statistique et économique standards peut entrainer des biais coûteux dans l'estimation de la dose optimale; d'autant plus qu'il y a un problème de pollution potentielle.

Une famille de modèles linéaire-plateau, à base de lignes droites se coupant est proposée pour reproduire l'effet de plateau. Les coefficients de régression sont calculables facilement, les interprétations économiques sont simples. On décrit des techniques d'interprétation économique et d'ajustement et d'estimation.

Pour des expériences à plusieurs éléments, on considère que l'expérience factorielle complète est le meilleur plan à la fois pour évaluer le modèle et estimer les applications optimales. L'information a priori peut fournir une base de décision: ainsi dans beaucoup de situations, seuls des plans d'étude de N ou NP sont nécessaires car les autres éléments nutritifs sont déjà à des niveaux corrects; on peut redistribuer la masse de matériel expérimental en ayant moins de facteurs mais plus de niveaux pour chaque.

On décrit deux plans couramment utilisés basés sur des informations préliminaires sur les doses optimales; le plan "un facteur à la fois" à le désavantage de ne pas fournir d'estimation de l'intéraction. Nous recommandons de concentrer plusieurs niveaux des traitements dans le voisinage de l'optimum subodoré. Puisque la phase "montante" du schéma de réponse est plus importante que la phase "plateau", elle doit recevoir plus d'attention.

REFERENCES

Anderson, R. L. and Nelson, L. A. [1971]. Some problems in the estimation of single nutrient response functions. *Bull. Int. Statist. Inst. 44*, Part 1, 203–22.

Baird, B. L. and Fitts, J. W. [1957]. An agronomic procedure involving the use of a central composite design for determining fertilizer response surfaces. In: *Economic and Technical Analysis of Fertilizer Innovations and Resource Use*, E. L. Baum, E. O. Heady, J. T. Pesek and C. G. Hildreth (Eds.), Iowa State College Press, Ames, Iowa.

Boyd, D. A. [1970]. Some recent ideas on fertilizer response curves. *Proc. Congress Int. Potash Inst., Berne, on the Role of Fertilization in the Intensification of Agricultural Production*, September 15–18 at Antibes.

Brown, William G. [1956]. Practical applications of fertilizer production functions. In: *Methodological procedures in the economic analysis of fertilizer use data*, E. L. Baum, Earl O. Heady and John Blackmore (Eds.), Iowa State College Press, Ames, Iowa.

Engelstad, O. P. and Parks, W. L. [1971]. Variability in optimum *N* rates for corn. *Agron. J. 63*, 21–3.

Fuller, W. A. [1969]. Grafted polynomials as approximating functions. *Aust. J. Agr. Econ. 13*, 35–46.

Waugh, D. L., Cate, Jr., R. B., and Nelson, L. A. [1973]. Discontinuous models for rapid correlation, interpretation, and utilization of soil analysis and fertilizer response data. *Technical Bulletin No. 7, International Soil Fertility Evaluation and Improvement Series*, North Carolina State University, Raleigh, North Carolina.

Wold, Svante. [1974]. Spline functions in data analysis. *Technometrics 16*, 1–11.

Received December 1974

Key Words: Response surface designs; Incomplete factorial designs; Linear-plateau model; Quadratic and square-root models; Economic optimum; Fertilizer experiments.