ON A MODEL USEFUL FOR APPROXIMATING FERTILIZER—RESPONSE RELATIONSHIPS

By
V.K. Gupta and A.K. Nigam
IASRI, New Delhi-12

(Received : March, 1981)

Summary

Second order response surfaces have been used widely by experimenters when the response is assumed to be related to the input-factors. This model has some disadvantages, though it is very simple in fitting. One disadvantage is the necessarily built-in symmetry about the modal value. An inverse-type polynomial, which is a second degree polynomial but which is asymmetric in form, is an alternative to the ordinary second order polynomial. It has been found that if the observations have a long tail to the right then the performance of this polynomial is better than the ordinary polynomial. For symmetric situations the two polynomials behave equally well. However, for negatively-skewed observations, the performance of ordinary polynomial is better than the performance of the inverse-type polynomial.

Introduction

The theory of regression analysis has been developed for situations where a variable is assumed to be related to one or more measurements made, usually on the same object. For example, many large scale surveys are conducted to gather data on biometrical characters of plants for making a forecast of the crop yield, on the basis of information on biometrical characters of plants, before the crop is actually harvested. Similarly a large number of experiments are conducted at various model agronomic experimental stations and on cultivator's fields to work out the suitable optimum doze of fertilizer for adoption by the cultivators under different agro-climatic conditions. The data gathered from such studies are summarized by fitting a suitable response surface. A typical model could be

$$y_u = \eta(x_{1u}, x_{2u}, \ldots\ldots\ldots; \beta_j) + e_u.$$

$\mathbf{y}=(y_1,\ldots\ldots, y_n)'$ is the output variable (or the dependent variable), $\mathbf{x}_1, \mathbf{x}_2,\ldots\ldots$ are the input factor levels (or independent variables), where $\mathbf{x}_i=(x_{i1}, x_{i2},\ldots\ldots,x_{in})'$, $\beta=(\beta_1, \beta_2,\ldots)'$ is the set of parameters unknown partly or wholly, e_u is a random error associated with the u-th observation. The e_u's are distributed independently and identically with zero mean and constant variance σ^2.

Determination of the form of η (.) is a serious and a difficult problem. A common practice is to choose for η (.) ordinary polynomials in which β are the co-efficients of the various terms in the polynomial. Such polynomials are extensively used as models in the analysis of response surface designs and are also implicitly used in analyses of factorial designs. The simplicity in the analysis using ordinary polynomials is, however, accompanied by some disadvantages inherent in the model (Nelder, [4]). The first degree polynomial is obviously unbounded. The second degree polynomial has a necessarily built-in symmetry about its modal value. There seems to be no reason to expect that the form of the response should be symmetrical about the modal value. An examination of the reports published by the Indian Agricultural Statistics Research Institute on the analysis of various model agronomic experiments and trials on cultivator's fields revealed that in majority of the cases the response curves were skewed positively though we observed a long tail to the left also in some cases.

Many models have been tried to circumvent the discrepancies, spelt above, in the ordinary polynomials, and these models have proved to be better than the ordinary polynomials. Abraham and Rao [1] used the five functions viz., Mitscherlich—Baule function, Generalized Cobb-Douglas function, Maskell Resistance formula, Quadratic response surface and Quadratic square-root transformation. In the case of quadratic functions, whether in variables or in their square roots, the response surfaces are linear functions of the constants and the fitting by least square method is therefore relatively easy. In case of other three surfaces the formulae are non-linear functions of the constants. Hence the fittings of the functions by least square method is difficult and involves heavy computations. Wold [5] demonstrated the usefulness of Spline functions over second degree polynomials. Here once again, the fitting of Spline-functions introduces computational difficulties. The inverse polynomial of Nelder is also difficult to fit because of the assumptions on error component. Anderson and Nelson [2] suggested a family of linear-plateau models for fitting fertilizer response data which exhibit a plateau effect. These models are quite useful but are limited in application to only those situations where the response exhibits a plateau.

In this paper, we propose a second degree polynomial which does not have a built-in symmetry and can be fitted easily using the theory of ordinary least squares. It has been noticed that the suggested polynomial fits better when the response curve is positively skewed as it has smaller residual sum of squares than the ordinary second-degree polynomial. Substantial work, however, needs to be done to investigate a model (polynomial or otherwise) for approximating negatively skewed response curves.

2. The Proposed Model

Nelder [4] introduced inverse polynomial response functions defined by

$$\frac{x_1 . x_2 \ldots\ldots x_k}{y} = \text{Polynomials in } (x_1, x_2, \ldots\ldots, x_k).$$

The two disadvantages inherent in ordinary polynomials are not prevalent in this model. It has been shown empirically by Nelder [4] that these polynomials prove better than the ordinary polynomials in the sense that these have smaller residual sum of squares as compared to the residual sum of squares of ordinary polynomials. This model, however, involves the following error structure :

$$E(\underline{e}) = \underline{o} \text{ and } D(\underline{e}) = \sigma^2 \text{ Diag } (y_1^2, \ldots, y_n^2)$$

The estimation of parameters, therefore, involves weighted least squares. Also, to compare the residual sum of square obtained from fitting ordinary polynomials, the ordinary polynomials have to be fitted subject to same error assumptions. Though the inverse polynomial response surfaces have provided encouraging results, the problem of estimation and then the comparison with other ordinary polynomials cannot be overlooked.

Draper and St. John [3] introduced the following linear model

$$y_u = \beta_o + \sum_i \beta_i \, x_{iu} + \sum_i \beta_{-i} \, x_{iu}^{-1} + e_u.$$

This model can be used with quite appreciable gains in experiments with mixtures, where $0<a\leqslant x_i\leqslant b<1$ and $\sum_{i=1}^{k} x_i=1$. However, there is a problem with this polynomial that it becomes discontinuous at $x=0$, when x takes values on the real axis. This problem does not arise when this model is used in experiments with mixtures. But in fitting response surfaces this problem will occur. We shall, therefore, prefer taking $x>E$, where E is an infinitesimally small positive quantity. In this paper we study the usefulness of this model in fitting response surfaces.

3. RESULTS

To study the usefulness of the proposed model the data on fertilizer-crop yield trials were used. The data were obtained for 14 experimental stations from a report published by Indian Agricultural Statistics Reseaich Institute, New Delhi. The following two models were tried :—

(*i*) Usual second degree polynomial

$$E(y)=\beta_0+\beta_1x_1+\beta_2x_1^2$$

and (*ii*) the proposed model

$$E(y)=\beta_0+\beta_1x_1+\beta_{-1}x_1^{-1}.$$

The results are summarized in Tables 1(a) and 1(b). The performance of these two models was more or less of the same nature except in five cases where the difference was quite appreciable. Out of these five cases, inverse type polynomial scored better than the ordinary polynomial in the sense that the former had smaller residual sum of squares than the latter. In the other two cases, the performance was otherwise. To obtain a reason for the performance, the gr aphs of actual yields were plotted. These are given in Figures 1—14. It has been observed that if the observations have long tail to the right, the performance of inverse-type polynomial is better than the performance of ordinary polynomial. In case the observations are symmetrical about the stationary value, the performance of ordinary polynomial remains unquestioned, though the performance of inverse-type polynomial is also more or less of the same order. However, if the observations exhibit a tail to the left, then the performance of ordinary polynomial is better than that of the inverse-type polynomial.

TABLE 1(a)

Model $E(y)=\beta_0+\beta_1 x_1+\beta_2 x_1^2$

			R^2
2611.15	9.1288 (1.7243)	−0.0848 (0.0138)	0.98
2538.64	10.5413 (5.4865)	−0.0477 (0.0438)	0.91
1218.09	37.0650 (14.1575)	−0.1631 (0.1131)	0.95
1020.49	21.5375 (29.4875)	−0.0847 (0.2355)	0.66
3854.00	−5.2000 (2.2753)	0.1156 (0.0182)	0.99
2520.14	22.0288 (13.2739)	−0.1573 (0.1060)	0.74
2080.88	252.5582 (68.5333)	−6.8861 (4.1074)	0.97
4785.59	35.0751 (39.7645)	−3.3438 (2.3832)	0.70
4670.71	38.6930 (48.8900)	−3.1964 (2.9301)	0.50
4166.11	288.0435 (108.3443)	−17.0715 (6.4934)	0.78
5808.54	−256.8222 (39.2797)	11.5983 (2.3541)	0.97
4383.37	−58.7107 (84.8202)	10.7991 (5.0835)	0.92
741.73	289.0545 (17.5456)	−11.9331 (1.0444)	0.99
682.33	283.1545 (22.9160)	−11.6831 (1.3734)	0.99

Figures in brackets are the standard errors.

TABLE 1(b)

Model $E(y)=\beta_0+\beta_1 x_1+\beta_{-1} x_1^{-1}$

			R^2
3040.66	−4.2000 (2.6703)	−0.0042 (0.0028)	0.75
2719.66	3.7000 (3.3632)	−0.0015 (0.0035)	0.85
1899.33	12.9875 (9.6057)	−0.0060 (0.0099)	0.90
1078.99	12.2000 (14.1162)	0.0011 (0.0146)	0.63
3206.99	13.6250 (1.7321)	0.0066 (0.0018)	0.99
3182.33	−1.2500 (9.1077)	−0.0059 (0.0094)	0.44
2576.00	119.7500 (35.8487)	−0.0046 (0.0051)	0.95
5153.00	−40.1250 (5.8315)	−0.0043 (0.0008)	0.97
5052.00	−35.7000 (8.5404)	−0.0045 (0.0012)	0.91
5799.00	−75.6502 (14.7568)	−0.0181 (0.0021)	0.98
4828.99	−20.5498 (26.2870)	0.0101 (0.0037)	0.93
3329.49	163.1002 (16.9134)	0.0118 (0.0024)	0.98
1682.50	31.5499 (32.7008)	−0.0093 (0.0046)	0.89
1578.50	52.6999 (35.3205)	−0.0087 (0.0050)	0.90

Figures in brackets are the standard errors.

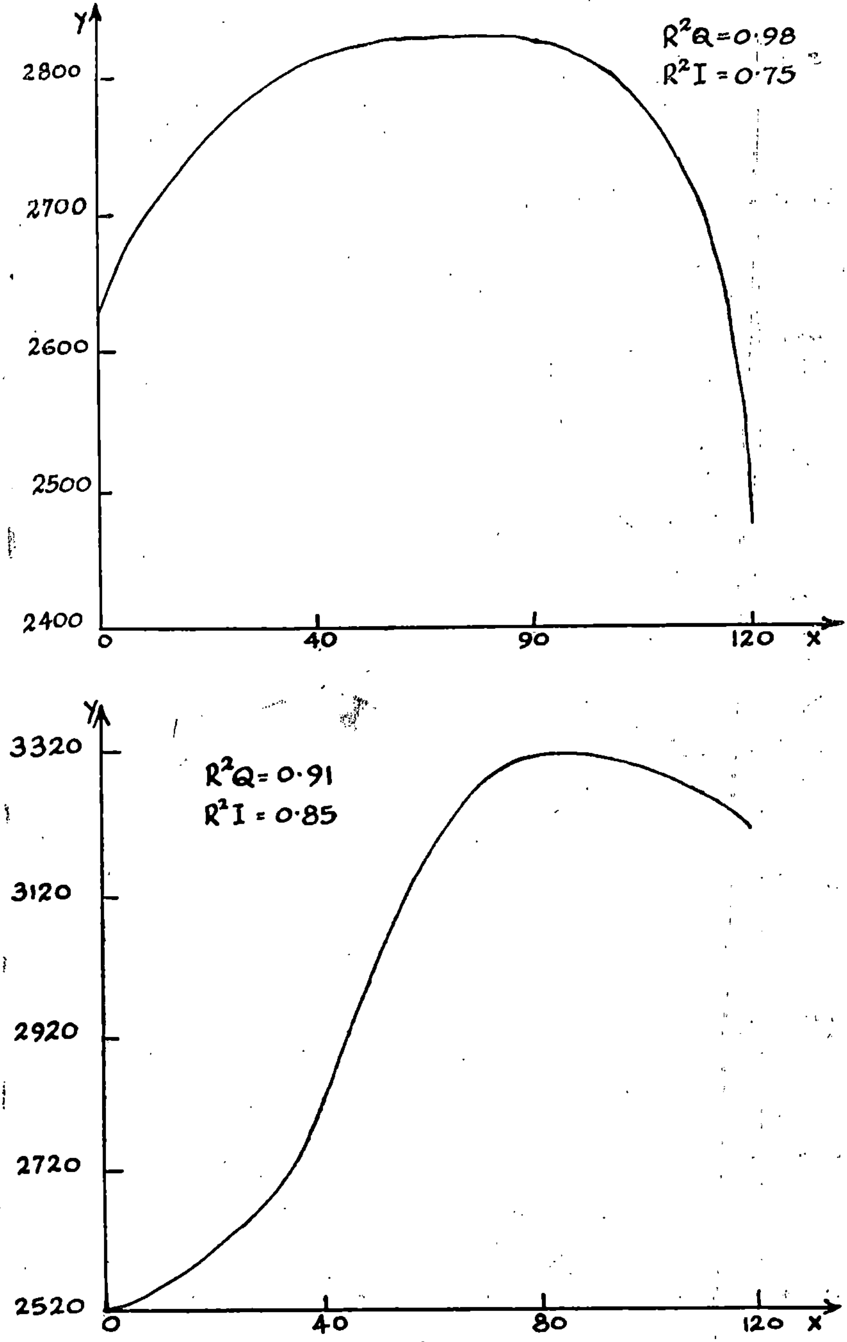

Y
2800
2700
2600
2500
2400
0
40
90
120
X
R²Q=0·98
R²I = 0·75
Y
3320
3120
2920
2720
2520
0
40
80
120
X
R²Q= 0·91
R²I = 0·85

Fig. 1 and 2

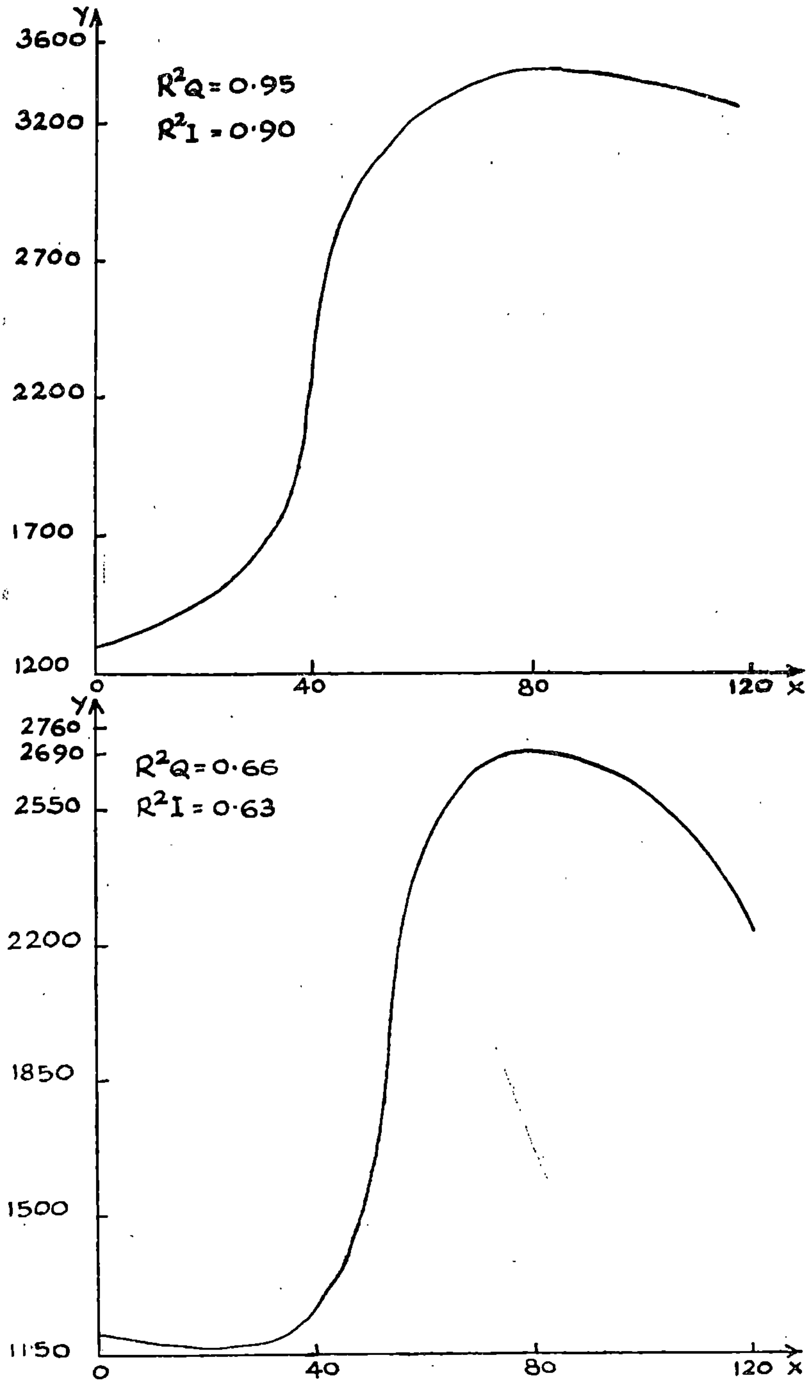
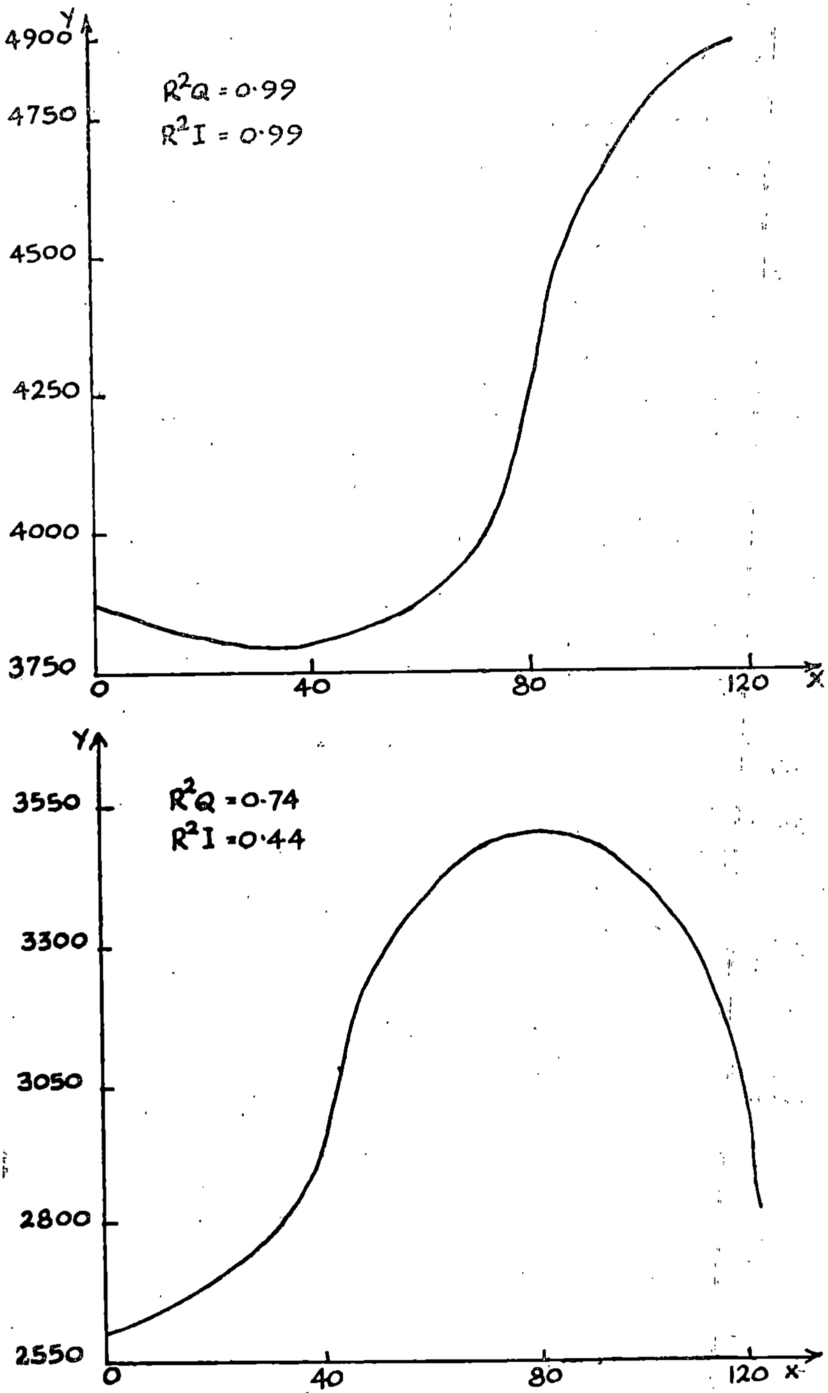

Y
3600
3200
2700
2200
1700
1200
$R^2Q = 0.95$
$R^2I = 0.90$
0
40
80
120 X
Y
2760
2690
2550
2200
1850
1500
1150
$R^2Q = 0.66$
$R^2I = 0.63$
0
40
80
120 X

Fig. 3 and 4



Y
4900
4750
4500
4250
4000
3750
R²Q = 0·99
R²I = 0·99
0
40
80
120
X
Y
3550
3300
3050
2800
2550
R²Q = 0·74
R²I = 0·44
0
40
80
120
X

Fig. 5 and 6

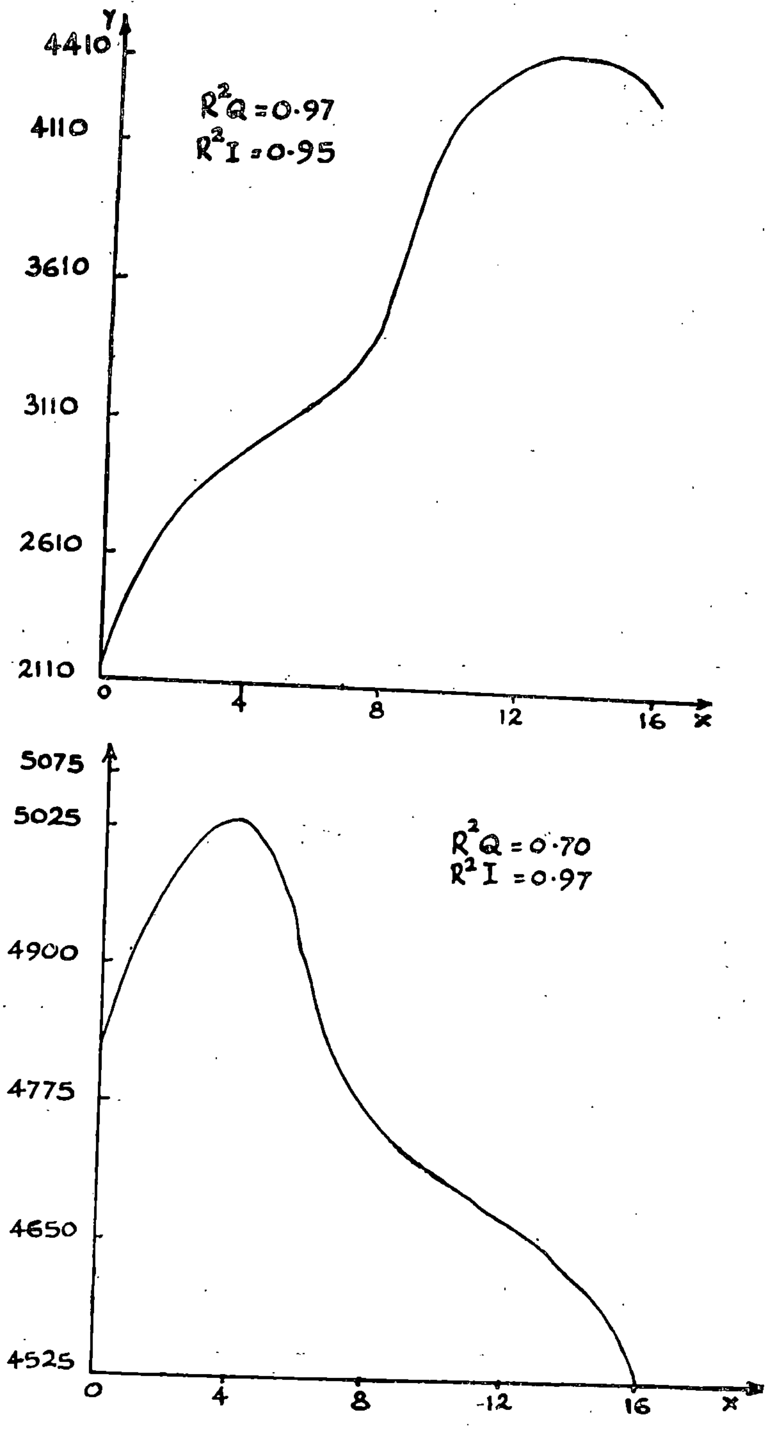

Y
4410
4110
3610
3110
2610
2110
$R^2Q = 0.97$
$R^2I = 0.95$
0
4
8
12
16
x
5075
5025
4900
4775
4650
4525
$R^2Q = 0.70$
$R^2I = 0.97$
0
4
8
12
16
x

Fig. 7 and 8

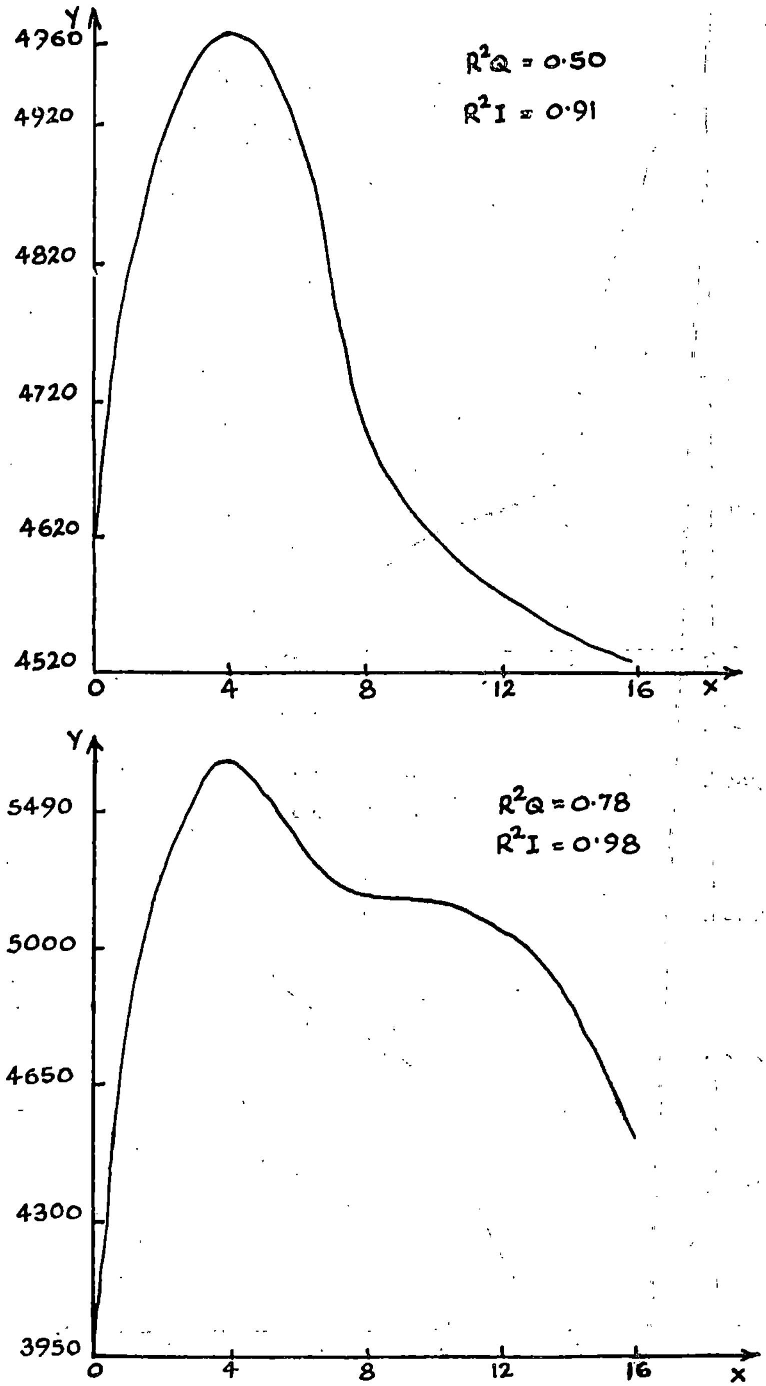

Y
4760
4920
4820
4720
4620
4520
0
4
8
12
16
X
R²Q = 0·50
R²I = 0·91
Y
5490
5000
4650
4300
3950
0
4
8
12
16
X
R²Q = 0·78
R²I = 0·98

Fig. 9 and 10

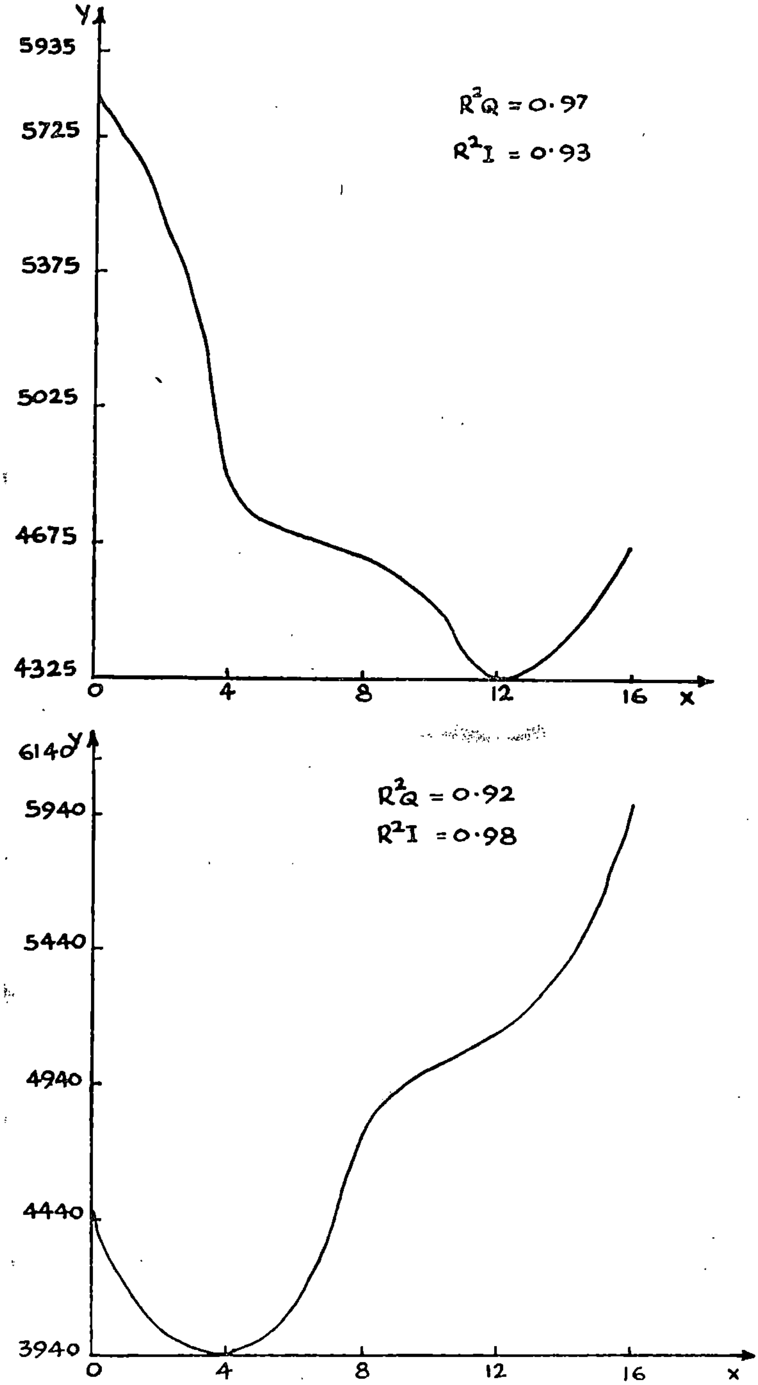

Y
5935
5725
5375
5025
4675
4325
0
4
8
12
16
X
R²Q = 0·97
R²I = 0·93
6140
Y
5940
5440
4940
4440
3940
0
4
8
12
16
X
R²Q = 0·92
R²I = 0·98

Fig. 11 and 12

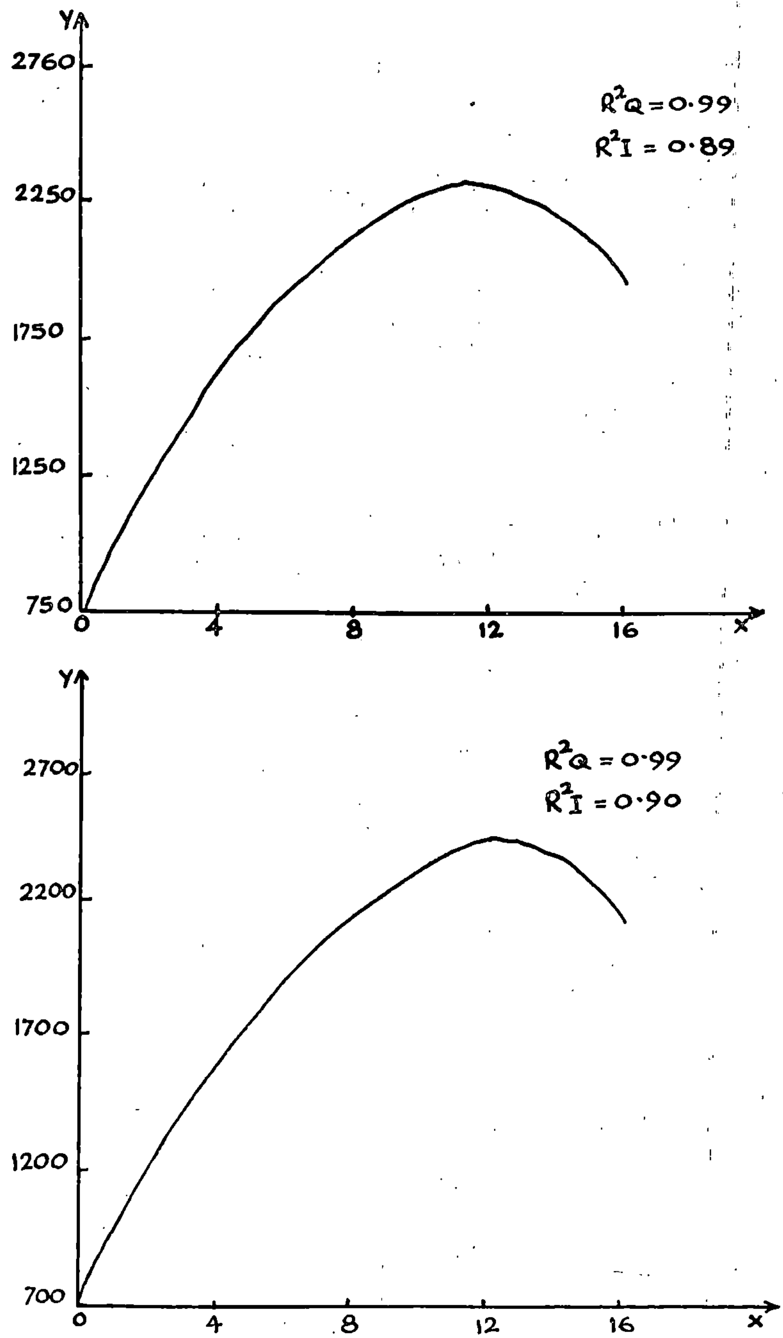

Y
2760
2250
1750
1250
750
0
4
8
12
16
X
$R^2Q = 0.99$
$R^2I = 0.89$
Y
2700
2200
1700
1200
700
0
4
8
12
16
X
$R^2Q = 0.99$
$R^2I = 0.90$

Fig. 13 and 14

ACKNOWLEDGEMENT

The authors are thankful to the referees for making useful suggestions.

REFERENCES

[1] Abraham, T.P. and Rao, V.Y. (1966) : An investigation on functional models for fertilizer-response surfaces. *Jour. Ind. Soc. Ag. Stat.*, **18**, 45-61.

[2] Anderson, R.L. and Nelson, L.A. (1975) : A family of models involving intersecting straight lines and concomitant experimental designs useful in evaluating response to fertilizer nutrients. *Biometrics*, **31**, 303-318.

[3] Draper, N.R. and St. John. R.C. (1977) : A mixture model with inverse term. *Technometrics*, **19**, 1, 37-46.

[4] Nelder, J.A. (1966) : Inverse polynomials, a useful group of multi-factor response functions. *Biometrics*, **22**, 128-141.

[5] Wold, S. (1974) : Spline functions in data analysis. *Technometrics*, **16**, 1-11.