ECONOMIC CONSIDERATIONS IN THE DESIGN AND ANALYSIS OF AGRICULTURAL EXPERIMENTS*

JOHN L. DILLON†

1. PREAMBLE

Roughly speaking, we might classify agricultural experiments into three types based on the question(s) which they aim to answer. These types are:

(i) "Where next?" experiments;
(ii) "Yes or No?" experiments; and
(iii) "How much?" experiments.

"Where next?" and "Yes or No?" experiments traditionally characterize "basic" research—they aim at developing and elucidating hypotheses about Nature in its purely physical garb. "Yes or No?" experiments have also been the traditional approach in our extension-oriented research—but that is a mistake.

"How much?" experiments aim at determining *best operating conditions* for a response process and are particularly pertinent to research for farm management extension purposes. In fact, "How much?" experiments generally yield information which subsumes all the information obtainable from "Where next?" and "Yes or No?" experiments, so it is a fair sign of scientific backwardness in our agricultural research that they have not been used more frequently to this end. But that is another story. Here, in the sure belief that better farming demands better experiments, I want to *outline* the role of "How much?" experiments in research for farm management extension. The word "outline" must be emphasized. Except for footnoting some of the more pertinent literature, no attempt will be made to note all the qualifications that might be raised for particular circumstances.

Traditionally in Australia, farm management extension recommendations (how much super?, how much concentrate?, etc.) have been based on "Yes or No?" experiments giving partial answers about technical efficiency (maximum physical yield) but generally ignoring economic efficiency (balance between marginal revenue and marginal cost). More and more, however, under the economic pressures that they face, farmers

---

* Paper presented at the N.S.W. Department of Agriculture's "Short-term School in Farm Management for Extension Officers" held at Grafton in October, 1965.

† Department of Farm Management, University of New England.

are concerned with information on *how much* of this or that input they might best use in particular production processes. A major way of providing such information is through "How much?" experiments.[1]

Before considering such experiments in detail, we might note the two (not unrelated) differences between "How much?" and "Yes or No?" experiments.

Firstly, while "Yes or No?" experiments are concerned with ascertaining if there are "significant" factor effects or "significant differences" between response to some few discrete levels of inputs, "How much?" experiments are concerned with estimating the mathematical relationship (factor-product relation or production function) between inputs and output. The difference is shown by a comparison of Figures 1 (a) and (b) below where output (Y) is measured on the vertical axis and input (X) on the horizontal axis.

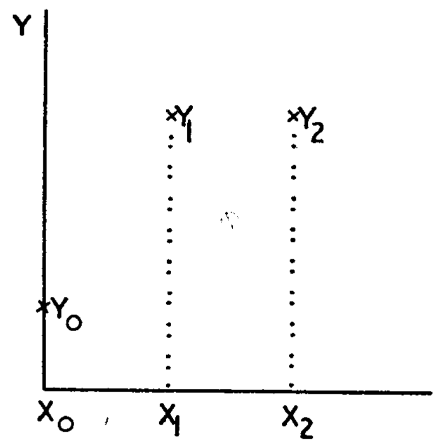

Figure 1 (a). "Yes or No?" experiment: Are there significant differences between Y_0, Y_1, Y_2?

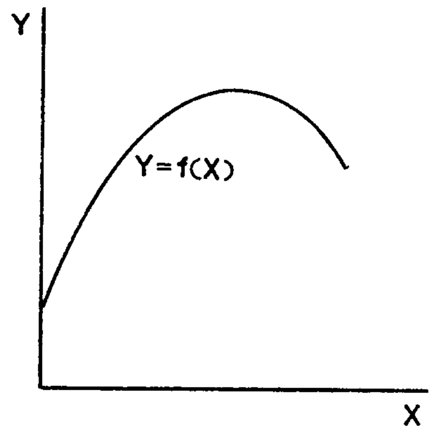

Figure 1 (b). "How much?" experiment: What is the mathematical form of the relation $Y = f(X)$?

Secondly, arising from this difference in aim, "Yes or No?" and "How much?" experiments differ in their statistical analysis. "Yes or No?" experiments, whether of truly discrete factors (e.g., tractors, cultivations) or of continuous factors (e.g., fertilizer, stocking rate) considered as discrete, typically imply *functional analysis of variance* as the mode of attack. In consequence substantial replication is needed not merely to increase the accuracy of the analysis of variance but to measure it. "How much?" experiments, whether of discrete or continuous factors, imply *least-squares regression analysis* (or some other method of function estimation). For such an analysis replication is not so essential.

---

[1] See A. G. Lloyd, "Agricultural Experiments and their Economic Significance", this *Review*, Vol. 26, No. 3 (September, 1958), pp. 185-209. Also of particular interest are three booklets in the series *O.E.C.D. Documentation in Food and Agriculture:* No. 50, "Inter-disciplinary Co-operation in Technical and Economic Agricultural Research" (Paris, 1962); No. 65, "Co-operation between Research in Agricultural Natural Sciences and Agricultural Economics" (Paris, 1964); and No. 71, "Co-operative Research to Improve Input/output Data in Cow Milk Production" (Paris, 1965).

Additional levels of a factor can substitute for replications of a particular level since additional observations help to locate the response curve more accurately whether they come from replications or extra factor levels.

Overall, the essential feature of "How much?" experiments is that they provide an estimate of the physical production function relating the quantity of output produced to the quantity of each input used in the response process.[2] In shorthand mathematical form such a production function is usually written—

$$Y = f(X_1, X_2, \ldots \ldots X_n);$$

indicating that Y (in some unspecified algebraic way) depends on the n input factors X_1, X_2, ... X_n. Estimates of such production functions are needed to provide the how-much answers corresponding to our best operating criterion of marginal cost equal to marginal revenue.[3] This criterion merely states that best operating conditions prevail when the last unit of input just pays for itself, i.e., denoting the last unit of the i-th input by $\triangle X_i$ and its contribution to output by $\triangle Y$, and the price per unit of input and output by p_i and p_y respectively, we must aim to have—

$$p_i \triangle X_i = p_y \triangle Y;$$

or, same thing—

$$\triangle Y / \triangle X_i = p_i / p_y.$$

If $p_i \triangle X_i$ is less than $p_y \triangle Y$, the last unit of input is more than paying for itself so that too little of it is being used. If $p_i \triangle X_i$ is greater than $p_y \triangle Y$, too much of the input is being used. These rules follow directly from the application of economic common sense to the law of diminishing returns. Note that while input and output prices are typically given from outside the production process, no matter how these prices may change, the physical production function remains relevant for conditions akin to those under which it was derived.

Against this background, let's look at the question of experimental design and analysis for "How much?" experiments. The aim of such experiments is to derive input quantities for best operating conditions defined by the above criterion or by the equivalent expression—

$$\delta Y / \delta X_i = p_i / p_y$$

obtained by using the differential calculus notation instead of the $\triangle$ notation. Thus we must be able to calculate the derivative $\delta Y / \delta X_i$, in turn implying that we have a production function which we can differentiate. To obtain such a function we generally have to use regression analysis, which in turn necessitates a certain type of experimental design so as to generate data in the required form. For this reason, it is logical to consider the mode of analysis before we consider the question of experimental design.

---

[2] Thorough development of the theory of physical production is to be found in R. Frisch, *Theory of Production* (Reidel: Dordrecht, 1965), Chs. 1 to 8. For specifically agricultural discussion, see R. D. Munson and J. P. Doll, "Economics of Fertilizer Use in Crop Production", *Advances in Agronomy,* Vol. 11 (1959), pp. 133-169.

[3] Should the response process be constrained in some way (for example, if only a limited sum were available to purchase inputs), the criterion for best operating conditions must be suitably modified. See Frisch, *op. cit.,* pp. 144-154.

2. REGRESSION ANALYSIS

The production function—

$$Y = f(X_1, X_2, \ldots\ldots X_n);$$

denotes a *response surface* involving $n + 1$ variables. For example, if Y is bushels of wheat per acre, X_1 is super per acre, and all the other factors X_2, X_3 . . . . X_n (i.e., nitrate fertilizer, number of cultivations, moisture, etc.) are held fixed, then the response surface would involve two variables (Y and X_1) and could be depicted as a curve in two-dimensional space, as in Figure 1 (b).

If two factors, say X_1 and X_2, are allowed to vary, and all the rest are held fixed, we have a surface in three-dimensional space, i.e., a hill, which we can depict geometrically either directly in three dimensions or indirectly from a bird's-eye view as a contour map. But if we vary three (or more) factors, we have a surface in four (or more) dimensions. We can't depict this geometrically, but we can do it algebraically (which of course we can also do for two or three dimensions). Thus for one, two and three variable factors respectively, we might have:[4]

$$\begin{aligned} Y = f(X_1) &= b_0 + b_1X_1 + b_{11}X_1^2; \\ Y = f(X_1, X_2) &= b_0 + b_1X_1 + b_2X_2 \\ &\quad + b_{11}X_1^2 + b_{22}X_2^2 + b_{12}X_1X_2; \\ Y = f(X_1, X_2, X_3) &= b_0 + b_1X_1 + b_2X_2 + b_3X_3 \\ &\quad + b_{11}X_1^2 + b_{22}X_2^2 + b_{33}X_3^2 \\ &\quad + b_{12}X_1X_2 + b_{23}X_2X_3 + b_{13}X_1X_3. \end{aligned}$$

For simplicity, we will concentrate on the three-dimensional surface, i.e., that relating to response from two variable input factors. All the implications of this case extend naturally to the case of more than two variable input factors.

In such terms our regression analysis simply implies having enough observations on Y and its corresponding levels of X_1 and X_2 so as to be able to calculate (by certain standard procedures which do not concern us here[5]) the numerical values of the coefficients b_1, b_2, b_{11}, b_{22} and b_{12} which arithmetically specify the two-factor response function shown above.

---

[4] Though the algebraic functions listed are all second-degree polynomials, other types of algebraic forms could be used as appropriate. However, since most functions can be satisfactorily approximated by a polynomial, the polynomial form is generally adequate. See E. O. Heady and J. L. Dillon, *Agricultural Production Functions* (Iowa State Univ. Press: Ames, 1961), pp. 203-209.

[5] See R. A. Bradley and J. S. Hunter, "Determination of Optimum Operating Conditions by Experimental Methods. Part I: Mathematics and Statistics fundamental to the Fitting of Response Surfaces; Part II: Models and Methods", *Industrial Quality Control*, Vol. 14, No. 1 (July, 1958), pp. 16-20; No. 6 (December, 1958), pp. 16-24; No. 7 (January, 1959), pp. 7-15; No. 8 (February, 1959), pp. 6-14. Also C. D. Throsby, "Fitting Production Functions to Experimental Data", this *Review*, Vol. 29, No. 3 (September, 1961), pp. 112-147.

[6] For crop-fertilizer experiments (x_{1r}, x_{2r}) is a treatment decided by (i.e., under the control of) the experimenter. In livestock feed experiments however, as discussed in Section 4 below, (x_{1r}, x_{2r}) is a treatment decided by the livestock—it is only indirectly under the control of the experimenter in so far as in the typical case he can dictate the ratio x_{1r}/x_{2r} but not the absolute values of x_1 and x_2 at particular points in time.

Hence, denoting factor levels by x's and observed response by y's, we need a set of data of the form shown in Table 1, where each doublet x_{1r}, x_{2r} denotes an experimental *treatment*[6] giving a response y_r. Should some or all treatments be replicated, these observations should be listed individually. They should not be averaged for regression analysis.

TABLE 1

Form of Data Needed for Regression Analysis

Observed	Factor Level	
Y	X_1	X_2
y_1	x_{11}	x_{21}
y_2	x_{12}	x_{22}
y_3	x_{13}	x_{23}
..	..	..
..	..	..
y_r	x_{1r}	x_{2r}
..	..	..
..	..	..
y_m	x_{1m}	x_{2m}

If we take a bird's-eye view of the response surface, then it is obviously sensible that the treatments should be well scattered over the surface. For example, consider the bird's-eye view of Figure 2 (a) with the response surface depicted by output contours (analogous to altitude contours as we move up the response hill from the origin). With nine treatments located at the (X_1, X_2) levels marked by the crosses, we would gain no information about the surface except on the line traced out by the treatments. In contrast, the layout in Figure 2 (b) with the same number of treatments would obviously tell us a lot more about the shape of the surface.

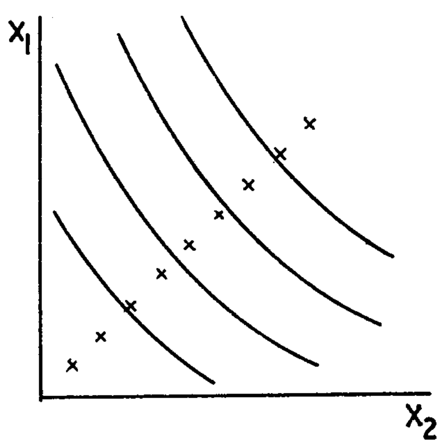

Figure 2 (a). Bird's-eye view of Y = f(X_1, X_2) with bad treatment layout.

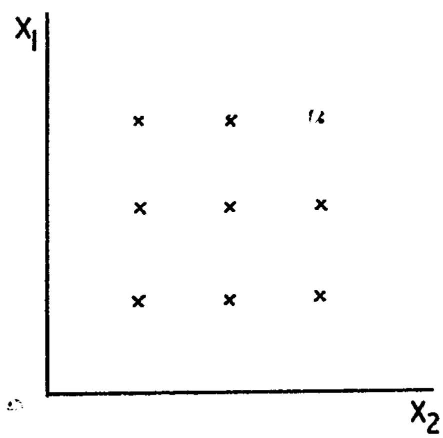

Figure 2 (b). Bird's-eye view of Y = f(X_1, X_2) with good treatment layout.

In general, therefore, the more purposively and systematically we cover the relevant surface region with treatments, the better the knowledge we can gain of the surface—and the better we can answer the question "How much?". In turn (for reasons which I will not elaborate), systematic surface coverage implies that relative to the traditional "Yes or No?" (functional analysis of variance) type of experiment, *experiments to ascertain best operating conditions imply:—*

(a) *more factors at more levels (with a minimum of three levels of each factor);*
(b) *a systematic arrangement of factor levels or ratios into treatments; and*
(c) *less emphasis on replication and the sacred cows of arbitrarily chosen significance levels that can bear only a coincidental relation to economic criteria for risky decisions.*

3. RESPONSE SURFACE DESIGNS

There are four types of experimental design that are particularly pertinent to response surface estimation.[7] These designs are the complete factorial, fractional factorial, central composite, and rotatable designs. We will give an example of each. For the reasons outlined in Section 4 below, while these designs are directly applicable to crop-fertilizer trials, they are not directly applicable to livestock-feed experiments.

COMPLETE FACTORIAL DESIGNS

The essence of these designs is that we choose our levels of each factor and then take all combinations of factors at all their levels as our treatments. Thus for two factors N and P, each taken at three levels, we would have the nine treatments (N, P) shown in Table 2.

TABLE 2

Treatments in a Three-level Factorial Design Involving Two Factors, N and P

Level of N	Level of P		
	p_1	p_2	p_3
n_1	n_1p_1	n_1p_2	n_1p_3
n_2	n_2p_1	n_2p_2	n_2p_3
n_3	n_3p_1	n_3p_2	n_3p_3

If the factor levels are equally spaced, we can code them for simplicity of analysis. Thus if the N and P levels of Table 2 each ranged over

[7] Standard references on experimental design for production function estimation are: W. G. Cochran and G. M. Cox, *Experimental Designs* (Wiley: New York, 1957), Chs. 1, 2 and 8A; Heady and Dillon, *op. cit.*, pp. 150-187; and N. L. Johnson and F. C. Leone, *Statistics and Experimental Design* (Wiley: New York, 1964), Sections 17.8-17.13.

values of 0, 100 and 200 lb per acre, we could code the factor levels as —1, 0, and + 1. Our nine treatments might then be listed as the *experimental points*—

(— 1, — 1)	(— 1, 0)	(— 1, 1)
(0, — 1)	(0, 0)	(0, 1)
(1, — 1)	(1, 0)	(1, 1).

These coded treatments can be diagrammatically represented by an "experiment map" as in Figure 3, where the coded treatment (0, 0) (corresponding to 100 lb per acre of both N and P) is the *centre of the design.* This design corresponds to that of Figure 2 (b) except that here we have coded the factor levels to give a new origin in terms of the experimental point (0, 0).

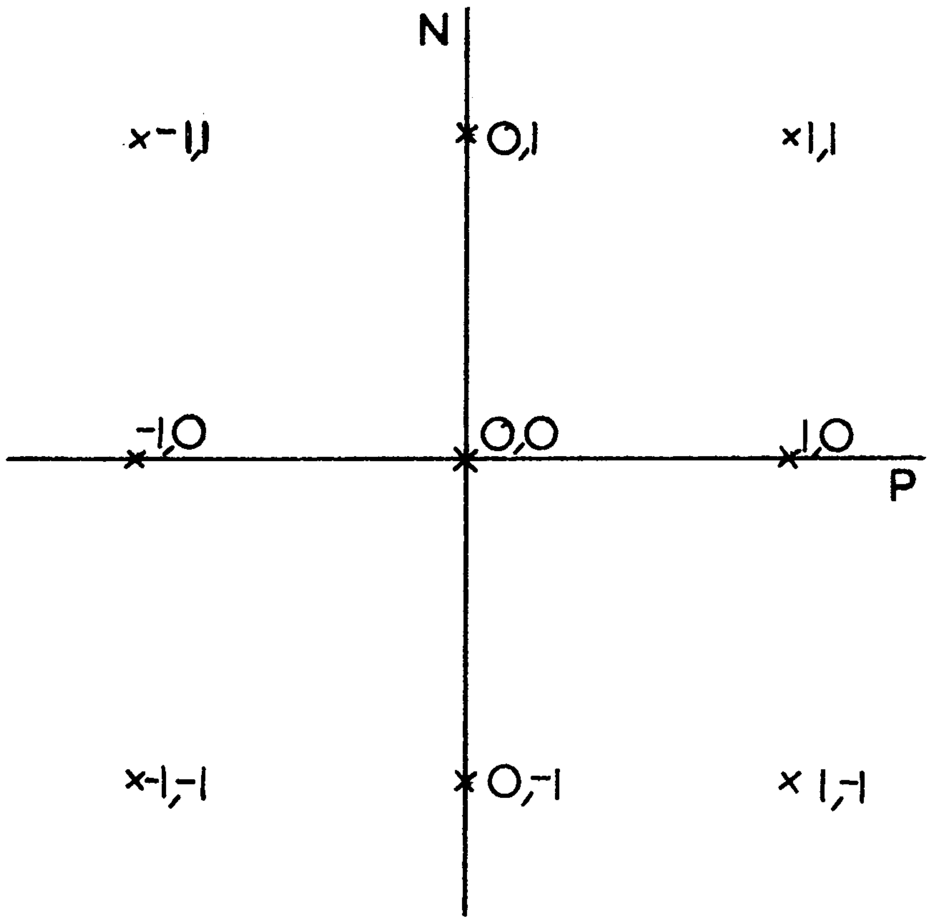

Figure 3. Experiment map corresponding to Figure 2 (b) with coded treatments.

In fact, complete factorials can be very inefficient for response surface estimation, especially if we allow for replication to gain an estimate of experimental error. Their scattering of treatments over the surface is so complete that, compared to the designs discussed below, they are wasteful of research resources if there are more than three factors to be studied.

FRACTIONAL FACTORIAL DESIGNS

These are obtained by taking only some fraction of the treatments from a complete factorial. For example, from the complete two-factor three-level factorial above, we might utilize the fractional factorial—

$$\begin{array}{ccc} (-1, \ 1) & & (1, \ 1) \\ & (0, \ 0) & \\ (-1, -1) & & (1, -1) \end{array}$$

or any other fraction that is reasonable and pertinent.

CENTRAL COMPOSITE DESIGNS

These designs are obtained by supplementing two-level factorials by additional experimental points or treatments arranged symmetrically about the centre of the factorial. For example, with two factors a central composite design would have the following treatments:—

$$\begin{array}{ccccc} & & ((0, \ \beta)) & & \\ & (-1, \ 1) & & (1, \ 1) & \\ ((-\beta, \ 0)) & & ((0, \ 0)) & & ((\beta, \ 0)) \\ & (-1, -1) & & (1, -1) & \\ & & ((0, -\beta)) & & \end{array}$$

where the extra treatments augmenting the basic factorial are in double parentheses.[8] The factor level $\pm \beta$ (coded value; $-1 > \beta > 1$) is at the choice of the experimenter. With only two factors, the central composite design gives no saving in treatment numbers over the complete factorial, but as the number of factors increases, the saving in number of treatments increases, as shown in Table 3. More importantly, compared to factorials, the central composite design gives a larger number of levels per factor. For example, the above composite design contains five levels of each factor compared with only three levels in the two-factor three-level factorial involving the same number of treatments. This is an important advantage since, as already stressed, the more levels of each factor the better for regression analysis.

TABLE 3

Number of Treatments per Replicate Needed for Factorial and Central Composite Designs

Design	No. of Factors				
	2	3	4	5	6
Factorial with 3 levels..	9	27	81	243	729
⅓ of 3-level factorial ..	..	9	27	81	243
Central composite ..	9	15	25	27*	43*

* Based on augmentation of a fractional two-level factorial.

[8] Note that while this design (as is typical of central composites) does not contain the zero treatment (i.e., $(-\beta, -\beta)$ if the coding is such that $-\beta$ corresponds to a factor level of zero), each factor does appear at zero level, though not both at the same time. If the full zero treatment is required, as might be the case for visual extension use, it could of course be added as an extra treatment.

ROTATABLE DESIGNS

In contrast to the factorial and central composite designs, rotatable designs have been developed specifically for response surface estimation. Like the central composite designs (which can be made rotatable by a suitable choice of experimental points), rotatable designs are particularly relevant for situations involving three or more input factors. Their essential features are: (a) that for given input units of measurement, they give estimates of response whose variance depends only on the distance from the centre of the design and not on the direction from the centre; and (b) that they enable a satisfactory estimate of experimental error to be obtained by replication of the central treatment so that, so long as there is no reason to expect that error will be directly related to the level of one or more treatments, it is not necessary to replicate the whole experiment, although if the experimenter wishes to, there is no reason why other experimental points in the design should not also be replicated.

An example of a three factor rotatable design suited to estimating a quadratic polynomial response function of the form—

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_{11}X_1^2 + b_{22}X_2^2 + b_{33}X_3^2 + b_{12}X_1X_2 + b_{23}X_2X_3 + b_{13}X_1X_3$$

is given in Table 4. All told, this design could involve a minimum of 16 treatments if we allowed but two observations on the central point to give an estimate of experimental error. Generally, however, some more replications of the central point would be desirable. Note that though the design may involve as few as 16 treatments, each of the three factors appears at *five* levels. In contrast, 27 treatments would be required for a *single* replicate of a *three*-level three-factor complete factorial.

TABLE 4

Coded Treatments for a Three Factor Second Order Rotatable Design

	X_1	X_2	X_3
Basic two-level factorial	—1	—1	—1
	1	—1	—1
	—1	1	—1
	1	1	—1
	—1	—1	1
	1	—1	1
	—1	1	1
	1	1	1
Treatments augmented to basic factorial	—1·682	0	0
	1·682	0	0
	0	—1·682	0
	0	1·682	0
	0	0	—1·682
	0	0	1·682
Replicates of central point	0	0	0
	..	..	..
	0	0	0

The design of Table 4 has been used for a 1965/66 U.N.E. Department of Farm Management maize-fertilizer trial on the Breeza Plain, Quirindi. Corresponding to the coded factor levels of Table 4, the treatments used in this trial are shown in Table 5 as active pounds of nitrogen, phosphorus and potassium.

TABLE 5

Levels of Nitrogen (N), Phosphorus (P) and Potassium (K) in a Maize-Fertilizer Trial using a Second Order Rotatable Design

Coded Value	—1·682	—1	0	1	1·682
Active lbs/acre of N ..	0	68·93	170·00	271·07	340·00
Active lbs/acre of P ..	0	48·18	118·80	189·42	237·60
Active lbs/acre of K ..	0	43·74	108·00	172·26	216·00

Specifications of many other rotatable designs are available in the literature. Suffice to note that for crop and pasture management trials of the "How much ?" type involving three or more input factors, *composite and rotatable designs are to be strongly recommended because—*

(a) *they are the most economic of available designs in terms of research resources; and*

(b) *they are specifically designed for response surface estimation.*

4. CROP VERSUS LIVESTOCK DESIGNS

The above discussion of experimental design is particularly appropriate for crop-management experiments. In such trials the experimenter can exercise control over the treatments studied. With livestock feeding experiments, however, there is a difficulty in that it may not be possible to control all aspects of the treatments.[9] To give a simple example, we may be able to specify what quantity of feed an animal can have, but not the time span over which it will eat this quantity of feed. Conversely, we may be able to specify the feed period but not the feed quantity to be consumed.

Diagrammatically, the difficulty shows up in that livestock feed treatments usually specify not points but a number of *ration lines* in the feed input plane, as exemplified in the pig-feeding trial depicted by Figure 4.

[9] See J. L. Dillon and H. T. Burley, "A Note on the Economics of Grazing and its Experimental Investigation", *Australian Journal of Agricultural Economics*, Vol. 5, No. 2 (December, 1961), pp. 123-132; and Heady and Dillon, *op. cit.*, pp. 240-250.

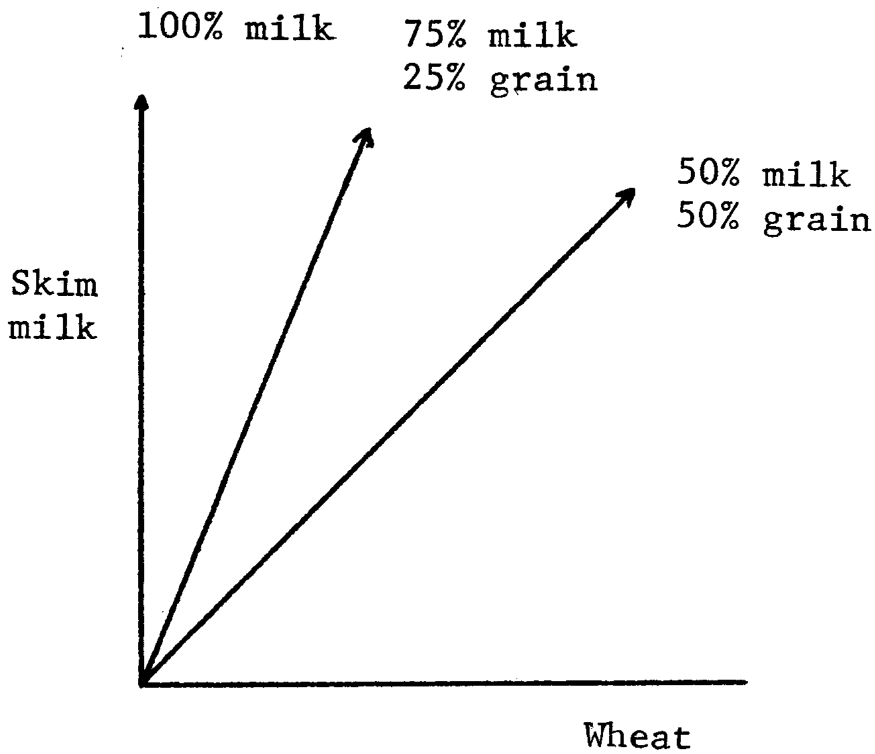

Figure 4. Ration lines followed in an ad libitum pig feeding trial.

As the experiment proceeds, over time the animals "move out along the ration lines". If observations are taken weekly, say for r weeks on feed consumed to date and liveweight, data for regression analysis is obtained as follows:—

Liveweight	Feed Consumption		Time
	Milk	Grain	
y_1	m_1	w_1	1
y_2	m_2	w_2	2
..	..	..	..
..	..	..	..
y_r	m_r	w_r	r

The treatments to be analysed are specified by the data quadruplets (y_t, m_t, w_t, t) which are best analysed on an individual pen basis rather than in terms of an average per replicate. In each treatment, only the level of time is under the experimenter's control. In consequence, in many livestock-feeding experiments the only feasible element of good design strategy for "How much?" experiments is to ensure that the ration lines give an adequate spread across the biologically feasible region of the response surface.

5. SOME COMMON CRITICISMS

Three criticisms are commonly made of response surface designs. Let's consider them in sequence.

Firstly, response surface designs may be more demanding of research resources than the traditional "Yes or No?" type of trial. Most often this criticism can be nullified by an appropriate choice of design; and even if this still implies relatively more research resources, the cost of these must be assessed against the extra information gained.

Secondly, results obtained from a particular experiment are only pertinent for the conditions under which the experiment was conducted. Though this is a severe criticism, we must not forget that it applies just as strongly to any other type of agricultural experiment. Moreover, it is of diminishing truth due to the development of statistical techniques—albeit demanding extra research resources—for meeting variability in response over space and time.[10]

Thirdly, there can be sizeable discrepancies between farm and experimental response.[11] For this reason, response surface experiments aimed at extension use should be carried out as far as possible under actual farm conditions.

6. CODA

1. Determination of best operating conditions implies knowledge of the response surface, i.e., of the continuous mathematical function relating inputs and output.

2. *Ipso facto*, experimental observations must generally be analysed by regression analysis and should therefore be generated in a form suited to regression analysis.

3. In consequence, so far as possible we should use designs oriented to response surface estimation—in particular, for crop-fertilizer trials, composite or factorial designs if there are two factors, and composite or rotatable designs if there are three or more factors; and for livestock-feed trials, a good spread of ration lines across the surface. Compared to the traditional approach involving much replication of a few factors at a few levels using functional analysis of variance for discrete effects, these response designs emphasize more levels per factor and much less wasteful use of research resources by unneeded replication.

4. Most importantly, in terms of economic analysis of experimental data, the economist is in the same boat as the statistician. Unless he participates in the planning stage of an experiment, there's a good chance he won't be able to properly apply economic principles to the analysis of the data. Farmers and others might then justifiably claim that research resources are being used inefficiently.

---

[10] Relative to variability over time, see W. A. Fuller, "Stochastic Fertilizer Production Functions for Continuous Corn", *Journal of Farm Economics*, Vol. 47, No. 1 (February, 1965), pp. 105-119. For spatial variability, see H. O. Hartley, "Experimental Designs for Estimating the Characteristics of Response Functions", *OECD Documentation in Food and Agriculture*, No. 65 (1964), pp. 163-176.

[11] See B. R. Davidson and B. R. Martin, "The Relationship between Yields on Farms and in Experiments," *Australian Journal of Agricultural Economics*, Vol. 9, No. 2 (December, 1965), pp. 129-140.