

CURRENT PROBLEMS IN DESIGN AND ANALYSIS OF FERTILIZER EXPERIMENTS

Peter Farnum

ABSTRACT

The statistical design and analysis of fertilizer experiments directed toward various scientific and management objectives are discussed. Early experimenters sought to determine whether fertilizers could increase tree growth, to identify what nutrient elements were limiting, and to quantify the order of magnitude of the response. Once the deficient elements had been identified, research studies were established to estimate responses to fertilizers applied at several dosage levels. The statistical designs used in these early studies were quite varied. Today scientists and managers are concerned with the development of fertilizer regimens. Generally the objective is to determine the economically optimal dosages and frequencies of application for stands of given densities and nutrient levels. The experimental designs and analytical techniques most suited to these current objectives are called response surface designs. Examples of response surface designs for both polynomial and nonlinear response functions are discussed for a number of experimental objectives. This methodology deserves more attention in fertilizer research.

INTRODUCTION

Originally I was asked to discuss current problems in the analysis of fertilizer experiments. It is my basic belief, however, and a central theme of this paper, that we cannot talk about analysis of experiments without also talking about objectives and design (Clutter 1968). This point is illustrated in Figure 1.

Figure 1. Interdependence of the major factors in a research plan.

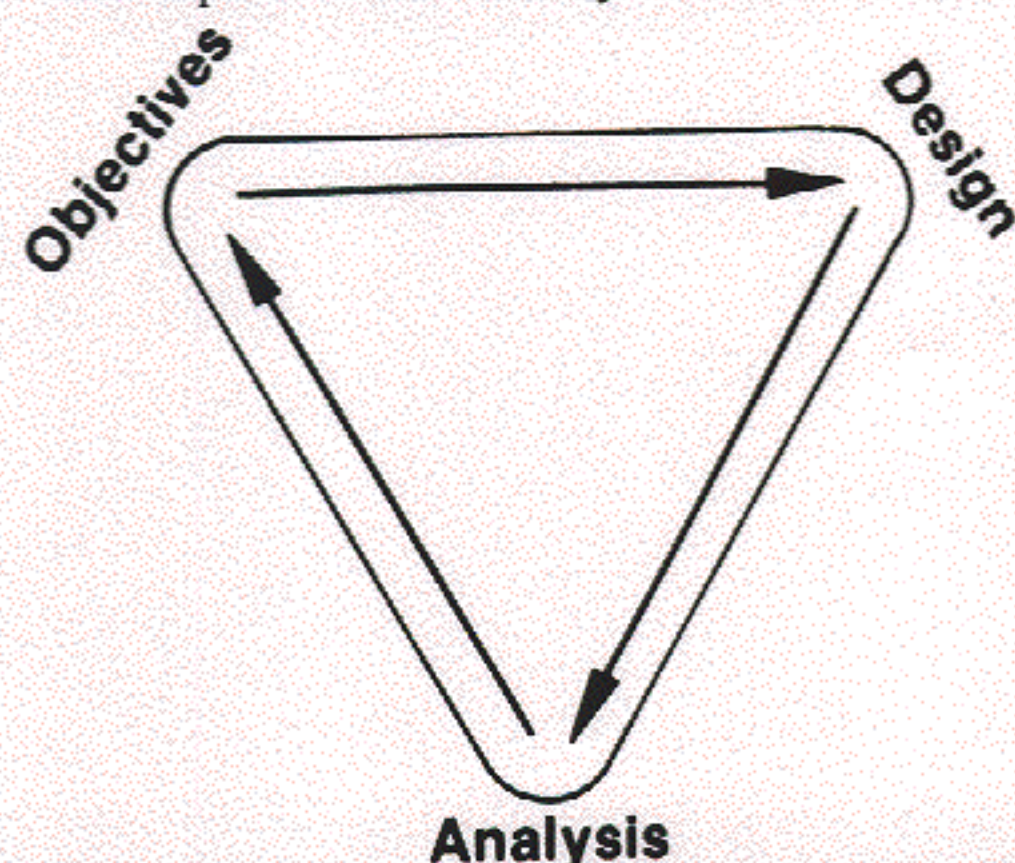


Figure 1. First, we set objectives for our experiment. Then, usually in consultation with a statistician, we arrive at an experimental design that will allow us to meet those objectives, and that determines, for the most part, the kind of analysis performed. After we complete this analysis, we evaluate whether or not we have successfully satisfied the original objectives.

The second basic theme of this paper is that there has been a logical evolution of experimental objectives, designs, and methods of analysis in fertilizer research. We have an objective and we design, carry out, and analyze an experiment to meet that objective. Most experiments raise as many questions as they answer, however, and in turn we design new experiments with the objective of answering these new questions. More often than not, new objectives require new designs and fresh methods of analysis—a continuing process.

Specifically, in fertilizer research much early work dealt with stands with obvious nutrient deficiencies, and the objective of these early experiments was to identify which elements were limiting and to quantify the order of magnitude of the potential response. After the limiting elements had been identified, later work sought to quantify this response more precisely for a series of levels of applied amendments. Finally, today's work is often concerned with development of optimal fertilizer regimens for those nutrients and stand conditions for which an economic opportunity has been identified.

This brings me to the third and major theme of this paper—that the designs and techniques of data analysis most applicable to developing optimal fertilizer regimens are given by what statisticians call response surface methodology (Box 1954, Box and Hunter 1956, Cochran and Cox 1957, Myers 1971, Mead and Pike 1975). After I present some historical examples of earlier designs, I will talk about this methodology and present some examples of the objectives, designs, and analyses for which it is suited.

EARLY RESEARCH

An example of a well executed early fertilizer research study is the work Heiberg and White published in 1950. They

reported results of a series of experiments in red pine plantations established on old fields in upper New York State. These plantations exhibited general chlorosis, low growth, short needles, and short needle retention. Heiberg and White hypothesized that some nutrient substance was lacking in the soil, and their objective was to identify the substance.

They installed six unreplicated plots with the treatments shown in Figure 2, and an untreated control; their analysis consisted of a similar graphical presentation. Clearly they observed a sizable growth response to potassium chloride and little or no response to the other treatments. Judging by these results, this was a remarkably successful experiment. No replication or sophisticated analysis was necessary because the size of the growth difference they were trying to detect was much greater than the natural plot-to-plot variation. This example illustrates the point that experimental designs and analyses evolve as experimental objectives change. While the experi-

ment was successful, unreplicated treatments and graphical analysis are not sufficient for today's questions.

After Heiberg and White had shown that K was limiting, they wanted to determine the most economical rate of K application. Hence they established experimental plots with four different application rates—the next logical step in the evolution of experimental objectives.

Similarly, in 1973 Miller and Pienaar reported on a study from the Wind River Experimental Forest in southwestern Washington. The objective of the study was to provide reliable data on volume response for a specific study area. The treatments were control, 140, 280, and 420 lb N/acre. There were three replicates of each treatment applied to 1/20th-acre measurement plots in a completely randomized experimental design. Miller and Pienaar analyzed their data using quadratic response functions fitted by regression (Figure 3).

This was an excellent piece of research from two stand-

Figure 2. Height growth response of red-pine to different nutrient amendments applied in 1943 (Heiberg and White 1950).

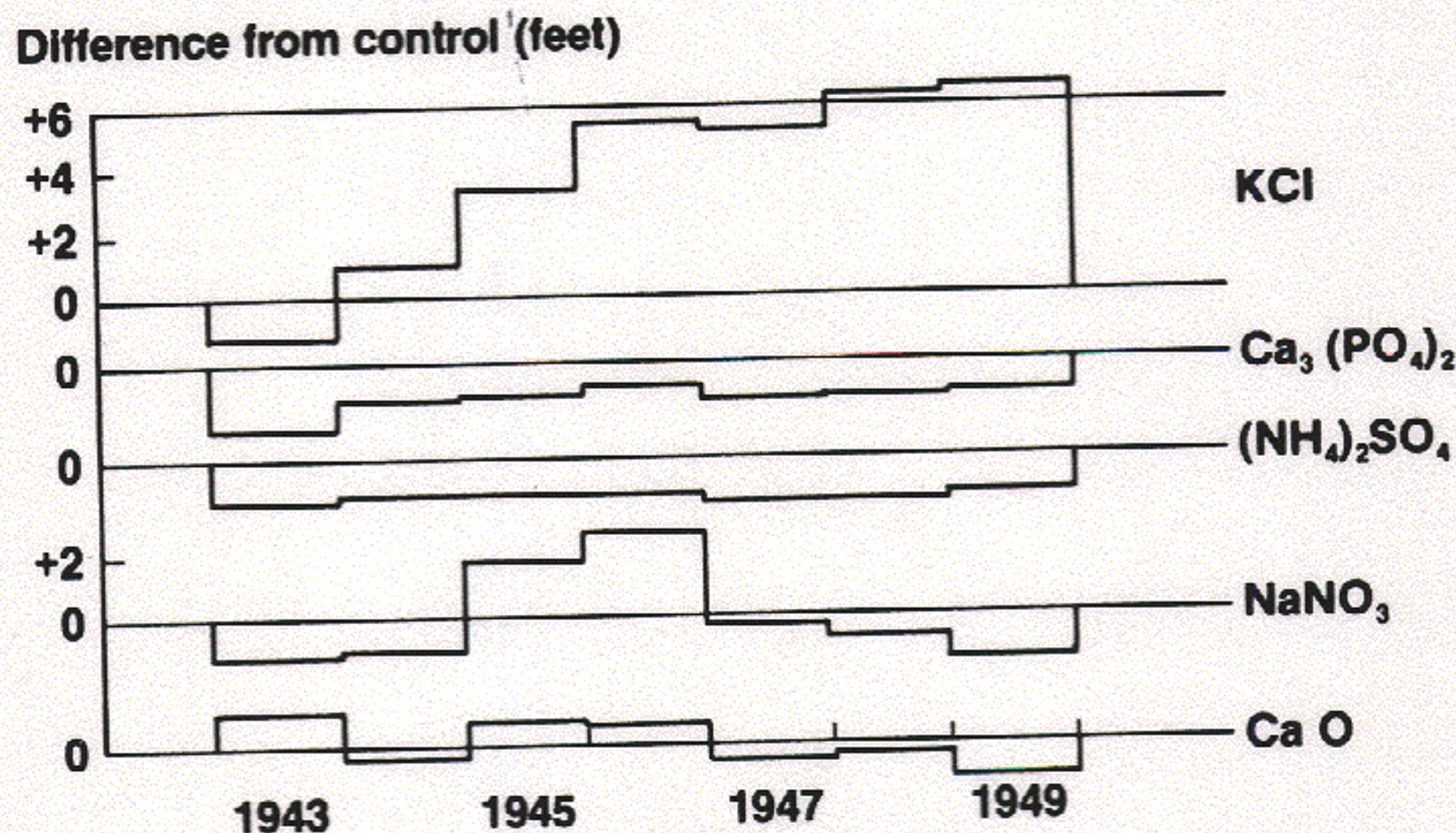
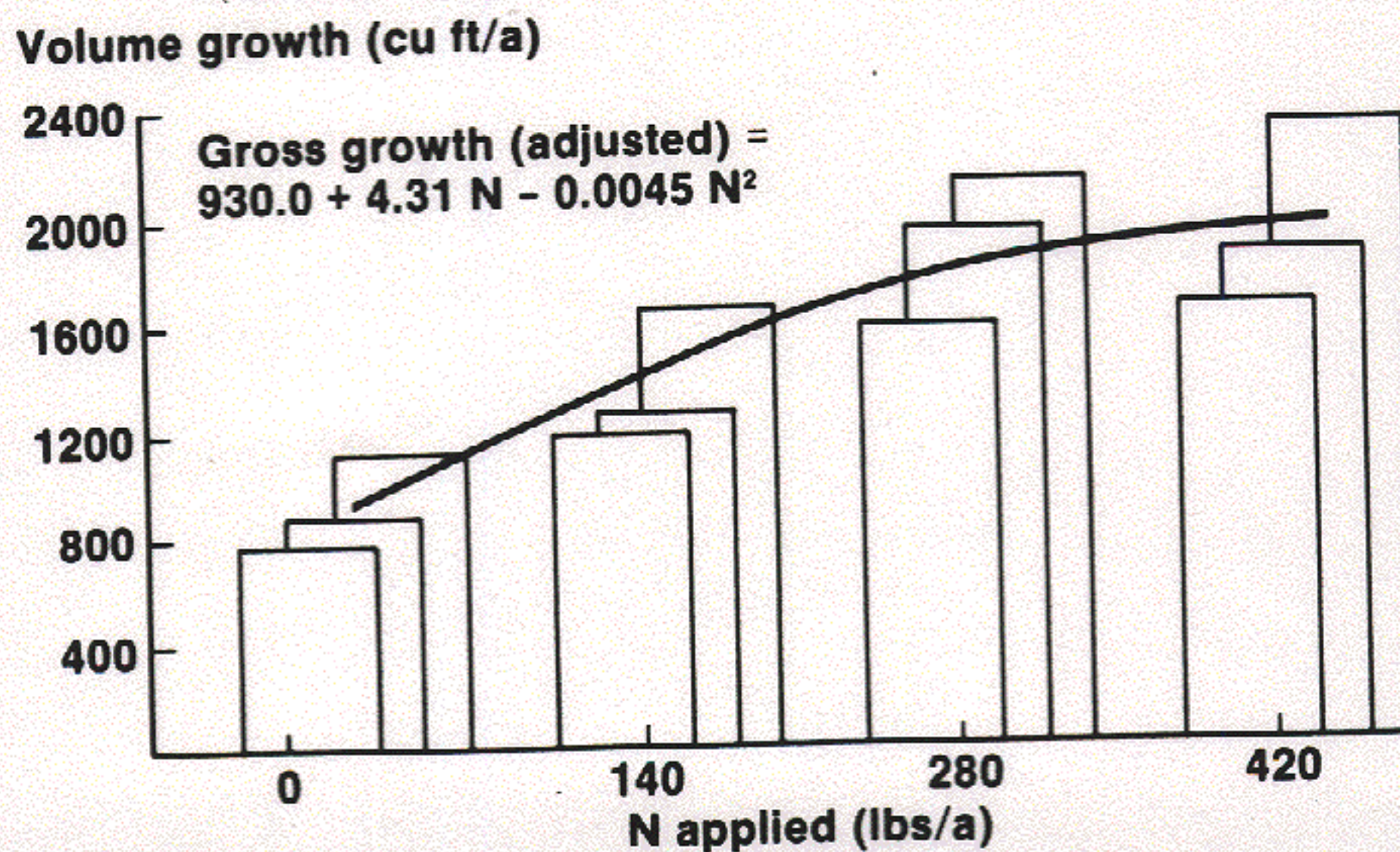


Figure 3. Volume growth of Douglas-fir at different levels of applied N (Miller and Pienaar 1973).



points. First, Miller and Pienaar avoided the temptation to use analysis of variance with a Duncan's multiple range test to look for significant differences between their treatments (Mead and Pike 1975). Treatments in the study were various levels of a specific quantitative factor, and the objective implicitly was to define the relation of response to these levels. The proper way to describe the relation is by fitting a mathematical function to the data, as they have done.

Second, they chose a quadratic response function, which is entirely appropriate for use with equally spaced data points over a limited range. They used this function to illustrate the declining marginal response that occurred as dosage increased over the range of their data. To have used these data to fit a more biological function (for example, a Mitscherlich equation, which rises to a plateau response) would have been risky and probably would have produced imprecise parameter estimates.

RESPONSE SURFACE METHODOLOGY

Miller and Pienaar's study raises important questions for us, both as biologists and as statisticians. First, now that we have a polynomial approximation of the relation between N dosage and volume response, we should ask what the true form of that function really is. Second, we have observed a specific pattern of response in a particular stand, but we must ask how the pattern will change with different levels of stand density or soil fertility. We need to determine what designs we should use to answer these questions.

The thesis of this paper is that the approach most suited to providing answers to questions in both these areas is what statisticians call response surface methodology. Response surface methodology is applicable for a biological study if the experimenter is dealing with continuous independent variables to occur at three or more levels of each treatment (Mead and Pike 1975).

Response surface methodology deals generally with two questions: (1) How do we select the treatments (that is, combinations of the independent variables) so as to fully explore the underlying response surface? (2) What analytical procedure do we use to summarize the data? Let me illustrate the response surface approach and demonstrate its advantages by example.

POLYNOMIAL FUNCTIONS

Consider the second question generated by the results of Miller and Pienaar's experiment, i.e., What are the effects and interactions of N dosage and stand density on the growth of young Douglas-fir plantations? The objective of our hypothetical experiment is to answer this questions.

Still following Miller and Pienaar, we recognize that we are at a rather early stage in the development of knowledge of this subject, so we will be satisfied with simply extending their

response polynomial to include the effects of stand density. We move from a quadratic function with N dosage as the independent variable to one with both N dosage and stand density as independent variables.

- Miller and Pienaar expressed growth as:

$$\text{growth} = b_0 + b_1 N + b_2 N^2$$

- An extension would be:

$$\text{growth} = b_0 + b_1 N + b_2 N^2 + b_3 D + b_4 D^2 + b_5 ND$$

There are shortcomings to the use of polynomial approximations, particularly for extrapolation beyond the observed data, but I believe use of these functions is defensible when we have little or no prior information on the shape of the response surface. The true response surface has some functional form and we may regard this polynomial as a truncated Taylor series expansion of that unknown function.

The problem now is to choose a desirable set of treatment combinations that will constitute our design. There are two designs we might employ (Cochran and Cox 1957). Table 1 shows the traditional 3^2 factorial design. If we have three replications of this design, we have 27 plots plus a number of no-N check plots. (I have used Miller and Pienaar's dosages just for consistency). As an alternative we can use the central-composite rotatable design (Table 2). The center point of this design is replicated five times. With two replications this

Table 1. The traditional 3^2 factorial design for studying the effects of fertilizer and density.

		Stand density 200 350 500 (trees/acre)		
Fertilizer (lb N/acre)	140	X	X	X
	280	X	X	X
	420	X	X	X

Table 2. The central-composite rotatable design for studying the effects of fertilizer and stand density.

		140	Stand density 200 350 500 560 (trees/acre)			
Fertilizer (lb N/acre)	80			X		
	140		X			X
				X X		
	280	X		X		X
				X X		
	420		X		X	
				X		

design requires 26 plots, or nearly the same number as the factorial.

The advantages of the second design over the first are that: (1) Five levels of each factor are included, instead of three. In this example I have used the extra levels to extend the range of the independent variables. (2) In the response surface design the variance of the predicted response depends only on the distance from the center of the design and not on the direction. This is desirable, since we have no *a priori* reason to want greater precision for certain dosage-density combinations than for others. Further, the prediction variances for the response surface design will be approximately equal within the center box (i.e., 200–500 trees/acre, 140–420 lb/acre). (3) The response surface design lends itself more readily to blocking than does the factorial design.

In the response surface design we may use incomplete blocks of seven plots each to increase within-block homogeneity, and thus precision. In the factorial design we must use a complete nine-plot replicate as a block unless we are willing to confound the dosage-density interaction with blocks and cut the precision of the estimate in half (Cochran and Cox 1957).

In presenting this example of a two-factor study, I have actually avoided using one of the strongest arguments in favor

of response surface designs. That is that they generally require fewer plots to study a given problem. If we compare the response surface designs with five levels of each factor with factorial designs with five levels (Table 3), then the response surface designs are clearly more efficient. With three or more factors, the differences in number of plots required are large enough to make the response surface approach feasible, while the factorial is not (it is clearly impossible to install a field study with 625 plots per replicate).

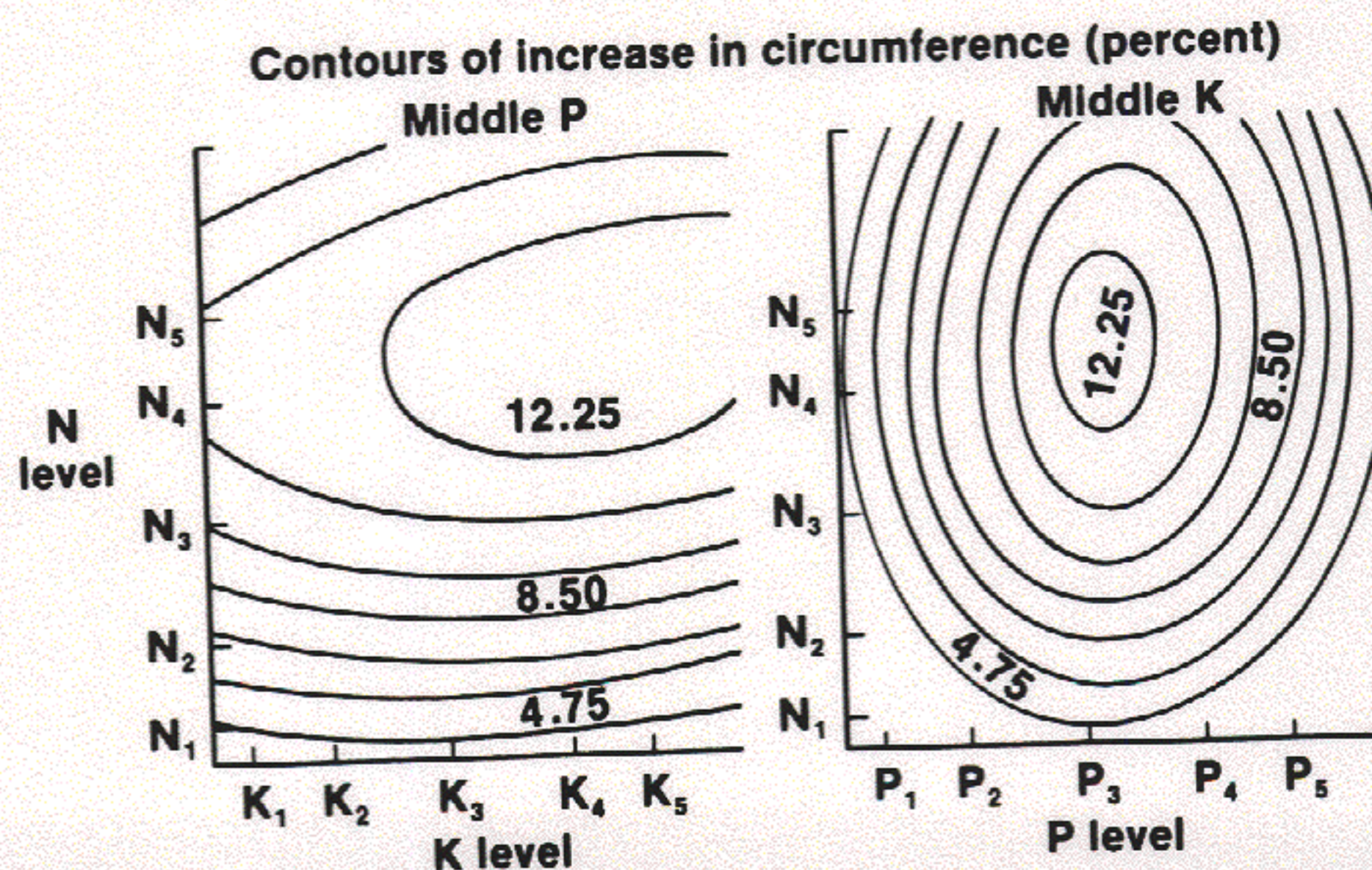
Before leaving the subject of response surface designs for polynomial response functions, I want to present results from an actual study reported in the literature. In 1978, Verma and Nijjar reported results of a study on the effect of N, P, and K fertilizers on growth of grapevines. They used a three-factor central-composite rotatable design with five levels of each nutrient in 20 treatment combinations. Their results are repeated here to illustrate the conceptual framework behind the response surface methodology.

Figure 4 shows the plotted contours of percentage increase in circumference as a function of (1) N and K dosage level at the middle level of P and (2) N and P dosage level at the middle level of K, respectively. Clearly, at the middle level of K there is a restricted combination of N and P that gives maximum response. If we move away from this combination in any direction, response falls off. In contrast, at the middle level of P there are rather broad combinations of N and K that give nearly maximum response. I think this is a productive way to conceptualize results of fertilizer trials and I question whether the same degree of understanding would have resulted if the experiment were done using a complete factorial design and analysis of variance.

Table 3. Number of plots required per replicate for factorial and response surface designs.

	Number of factors		
	2	3	4
	plots/replicate (no)		
5 ⁿ factorial	25	125	625
Response surface	13	20	31

Figure 4. Contours of percentage increase in circumference due to application of N, P, and K (after Verma and Nijjar 1978).



NONLINEAR FUNCTIONS

If the response surface methodology were restricted to the polynomial response functions I have used so far, its utility would be severely limited. To illustrate the response surface approach for nonlinear functions I return to the other question we generated after looking at Miller and Pienaar's results. This question has to do with the shape of the response surface when we apply N to Douglas-fir stands.

Evidence is accumulating that, over the range of interest, response increases rather quickly and then quickly levels out into a plateau (Figure 5). We can ask (1) where the knee of the plateau occurs, for that determines the optimal dosage; (2) what the level of the asymptote is, for that determines the attainable response; and (3) whether the response is sigmoid at very low dosages. We can ask these questions for each site class, and finally we can ask how to design an experiment to find the answers.

In this conceptualization the knee occurs at higher dosage levels on lower sites, which seems biologically reasonable. The level of the asymptote also changes by site class. Here I estimated the asymptotes from the 4-yr Regional Forest Nutrition Research Project 400-lb/acre results (Turnbull and Peterson 1976). Finally, I have hypothesized a region of increasing marginal response at low dosages.

How do we design an experiment to test the hypothesis summarized by this figure and to estimate what the true response surface looks like? Basically the steps are as follows.

First we choose a function that incorporates the hypothesized behavior. Here I propose moving from Miller and Pienaar's quadratic function to a Mitscherlich-type response curve. This function is capable of describing the hypothesized behavior. The parameter b_1 gives the asymptote, b_2 governs

the rate of approach to the asymptote, and b_3 controls the sigmoid behavior.

• Miller and Pienaar expressed growth as:
growth = $b_0 + b_1 N + b_2 N^2$

• An extension would be:
growth = $b_0 + b_1 (1 - \exp(-b_2 \cdot N))^{b_3}$

Next we determine the dosage levels that lead to the most precise estimates of these parameters. The specific mathematical approach used is to minimize the determinant of the $X'X$ matrix. This is equivalent to minimizing the hypervolume of the joint confidence region of the parameters. It may be shown mathematically that the minimum volume is attained when the number of design points equals the number of unknown parameters (Beck and Arnold 1977). In the particular case of this function, we must put constraints on the region of interest because the asymptote b_1 is estimated most precisely by an infinite dosage and the sigmoid parameter is estimated most precisely by a dosage that approaches zero. If we take the minimum and maximum dosages we wish to study to be 50 and 500 lb/acre, respectively, we find the middle dosage to be 225 lb/acre. This gives us our basic design points.

Incidentally, to make these calculations I used initial parameter estimates taken from Weyerhaeuser's empirical fertilizer trials, with design points as shown in Table 4. While difference in designs does not appear great, the determinant of the $X'X$ matrix in the response surface design is 1.7 times that in the equally spaced design.

Finally, we add extra points for lack of fit. With three parameters and three data points, we have no way of testing whether or not our chosen function actually fits the data. I therefore recommend installation of at least two additional treatment levels (perhaps replicated a few number of times) to test for lack of fit. This design would be repeated several times for each site class. The chosen equation should be fitted separately to each installation, tested for lack of fit, and then compared between every installation and every site class.

Thus, by applying results from response surface methodology, we have developed a sound experiment. The methodol-

Figure 5. Hypothetical volume response curves to N applied in Douglas-fir stands for three different site classes.

Volume response (cu ft/a)

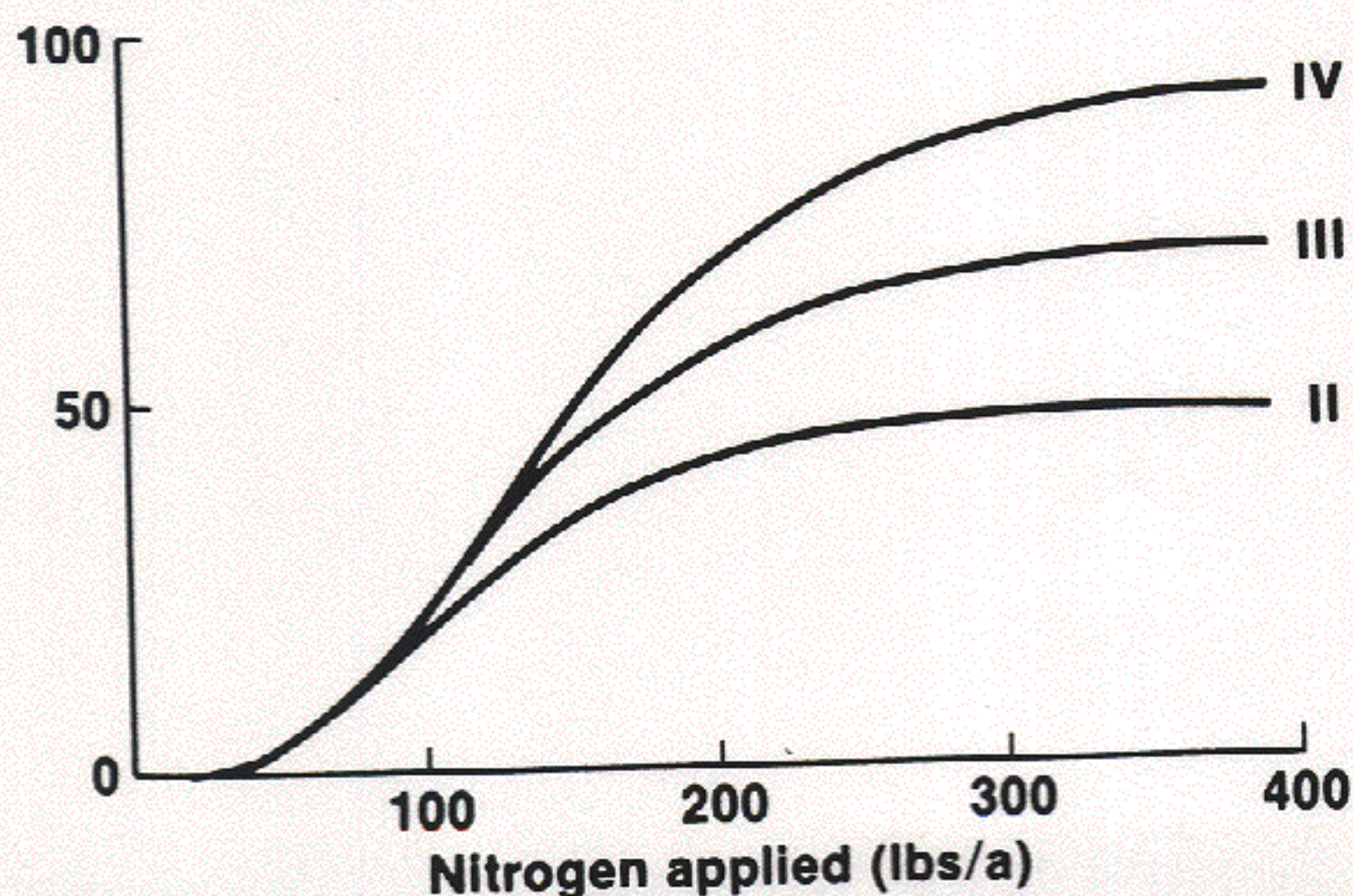


Table 4. Comparison of design points and relative variances for an equally spaced design (Weyerhaeuser) and a response surface design.

	Treatment level (N lb/acre)			Relative variance
Response surface	50	225	500	0.59
Weyerhaeuser 1969	100	300	500	1.00

ogy helped us define our objectives, determine our experimental design, and plan our analysis. All three aspects fit neatly together. We could not have arrived at this design without the discipline of response surface methodology.

CONCLUSIONS

Two points should be clear here. First, as our knowledge of forest fertilization increases, our experimental objectives, designs, and techniques of analysis must change. Designs and techniques that were appropriate for yesterday's questions are not adequate today. Second, a response surface methodology is available in the statistical literature to help us solve the current problems in design and analysis of fertilizer experiments. This methodology deserves more attention in fertilizer research.

LITERATURE CITED

- Beck, J. V., and K. J. Arnold.
1977. Parameter estimation in engineering and science. 501 p. John Wiley and Sons, New York.
- Box, G. E. P.
1954. The exploration and exploitation of response surfaces: Some general considerations and examples. *Biometrics* 10(1):16-60.
- Box, G. E. P., and J. S. Hunter.
1956. Multi-factor experimental designs for exploring response surfaces. *Ann. Math. Stat.* (1956):195-241.
- Clutter, J. L.
1968. Design and analysis of forest fertilization experiments. *IN* Forest fertilization—Theory and practice. G. W. Bengtson, ed., p. 281-288. Tennessee Valley Authority, National Fertilizer Development Center, Muscle Shoals, AL.
- Cochran, W. G., and G. M. Cox.
1957. Experimental designs. 617 p. John Wiley and Sons, New York.
- Heiberg, S. O., and D. P. White.
1950. Potassium deficiency of reforested pine and spruce stands in northern New York. *Proc. Soil Sci. Soc.* 1950:369-376.
- Mead, R., and D. J. Pike.
1975. A review of response surface methodology from a biometric viewpoint. *Biometrics* 31:803-851.
- Miller, R. E., and L. V. Pienaar.
1973. Seven-year response of 35-year-old Douglas-fir to nitrogen fertilizer. USDA For. Serv. Res. Pap. PNW-165. 24 p. Pac. Northwest For. and Range Exp. Stn., Portland, OR.
- Myers, R. H.
1971. Response surface methodology. Allyn & Bacon, Boston.
- Turnbull, K. J., and C. E. Peterson.
1976. Analysis of Douglas-fir growth response to nitrogenous fertilizer, Part 1: regional trends. For Res. Tech. Note 13. 15 p. College of Forest Resources, Univ. Washington, Seattle.
- Verma, H. S., and G. S. Nijjar.
1978. Response surface studies of the effects of N, P, and K fertilizers on vine growth, yield and fruit quality. *J. Hort. Sci.* 53:163-166.