## By

ACKELLO-OGUTU.
B. Sc. (University of Nairobi) 1975 M.Sc. (University of Nairobi) 1978

## DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of DOCTOR OF PHILOSOPHY in

Agricultural Economics in the GRADUATE DIVISION of the

## UNIVERSITY OF CALIFORNIA

DAVIS
Approved:


Committee in Charge

Deposited in the University Library $\qquad$

## Abstract

The study is motivated by the rapidly escalating energy prices, the accumulation of the major nutrients in many prime agricultural soils and the need to recognize the dynamic nature of the crop fertilization problem. The major objective of the study is to statistically test the appropriateness of alternative specifications of the relationship between crop response and factors of production. The hypothesis of nutrient nonsubstitutability is given special attention.

A model of corn response is formulated as a simultaneous equations system with response, soil phosphorus and soil potassium as the endogenous variables. The system is assumed to be recursive and the equations are therefore estimated separately as opposed to applying simultaneous equations techniques.

The equations estimating the soil level of the two nutrients are referred to as the carryover functions and are given a distributed lag model interpretation after introducing lagged values of the endogenous variables as exogenous variables.

The appropriateness of two kinds of specification of corn response functions is examined. The first kind is a polynomial approximation which has frequently been applied by many agricultural economists and soil scientists and by now considered as the conventional form for depicting crop response surfaces. The second kind is a nutrient non-substitution formulation.

The non-substitution model generalizes Liebig's "law of the minimum." It gives the expected corn yield (in the two nutrient case examined in
the study) as $\hat{A} \operatorname{Min}\left[f_{p}(P T), f_{k}(K T)\right]$ where $\hat{A}$ is the estimated yield maximum and the arguments in $f_{p}(\cdot)$ and $f_{k}(\cdot)$ representing estimated total availability of phosphorus ( $P$ ) and potassium (K), respectively. The indi-vidual-response functions $f_{p}(\cdot)$ and $f_{k}(\cdot)$ are approximated by linear splines and estimated by applying a nonlinear mathematical programming technique to a nonlinearly constrained problem. The estimation procedure, unlike those used hitherto in fitting the "law of the minimum," requires no prior sorting out of data in order to separate those for which either $P$ or $K$ is limiting growth.

Two polynomial formulations are estimated: the square root and the quadratic forms. Both indicate a general lack of response to $P$ thus pointing to an oversupply of this nutrient in the experimental plots. These results are verified by those of the non-substitution model which indicates that the maximum corn yield could be obtained by total nutrient supplies no more than $165 \mathrm{~kg} / \mathrm{ha}$ for both P and K .

Three non-nested hypothesis testing procedures are employed in testing the nutrient non-substitution hypothesis. The statistics used are Cox-Pesaran Statistic, C-statistic and linearized Cox-Pesaran statistic. The three test statistics give consistent results for small samples. For larger samples, the C-statistic gives results conflicting with those of the Cox-Pesaran statistic. The linearized Cox-Pexaran statistic is not found to be useful.

On the basis of the Cox-Pesaran statistic, it is concluded that the two polynomial forms cannot be supported against the nutrient non-substitution hypothesis given the data analyzed in the study. The nutrient non-substitution model is therefore proposed as a generalized tool which is not only biologically more appealing but is also statistically more
appropriate than the conventional approximations currently employed in crop response analysis.

AC'KELLO-OGUTU
Davis, California
29th March, 1982

## Acknowledgments

Many people have contributed either directly or indirectly to the assembly of what is expressed here in black and white. The broad variety of my academic and non-academic interactions at Davis have been immensely rewarding. An attempt to enumerate everyone within the limited space provided here will therefore not be without injustice to many. However, I shall take a crack at a list of those whose contributions to my dissertation are most vividly remembered and I hope my memory will not let me down.

My first acknowledgment goes to the chairman of my dissertation committee, Professor Q. Paris, for having conceived the original problem investigated in this study and for assisting me with total devotion and undiminishing interest. The useful comments of the other members of my dissertation committee Professors R. Green and W. A. Williams are also sincerely appreciated.

If anybody knows the computational trips I went through, then it must be Ron Huff. Data manipulation would have been such a big hastle without Ron's invaluable programming services which at times were called upon at rather odd hours.

I am indebted to Professor S. A. Barber for providing a much needed set of data, and I hope that my analytical results will convey what is intrinsically in the data and not the other way around.

It is often hard to magically convert frustration and bitterness to smiles when the core of the offending matter is pecuniary in nature. I must therefore thank Ms. Linda Childress and her competent staff at the Davis International Visitors Office for having been empathetic to
my problems, most of which were of the kind just mentioned.
I must thank Ms. Barbara Huffine for her unfaltering devotion to efficient and fast service on a broad range of aspects and also for her constant words of encouragement from as far back as when I first joined the Agricultural Economics Department at Davis.

My appreciation also goes to those fellow Kenyan students at Davis who made my life outside the domain of academics very pleasant.

I am also grateful to Professor Adolf Weber for getting me interested in graduate studies in agricultural economics and to the University of Nairobi and U.S.A.I.D. for making the scholarship a reality.

Last, but by no means least, I thank my family for their patience during my protracted student life. Special thanks go to my mother for her constant prayers for my success and for waiting so long for what, in her own mind, is intangible and the derived benefits anything but guaranteed in a world full of uncertainties.
CHAPTER 1: BACKGROUND ..... 1
CHAPTER 2: PROBLEM DEFINITION ..... 9
2.1 Formulation of Crop Response Models ..... 9
2.1.1 The need for a careful characterization of the response surface ..... 9
2.1.2 Polynomial approximations of crop response models ..... 11
2.1.3 Objections to polynomial approximations ..... 15
2.1.4 Incorporating site and weather characteristics in crop response functions ..... 25
2.2 Testing Model Specification ..... 26
CHAPTER 3: METHODOLOGY AND ESTIMATION TECHNIQUES ..... 33
3.1 The Nutrient Non-substitution Crop Response Model ..... 33
3.2 Formulation of the Nutrient Carryover Functions ..... 35
3.3 Estimation of the Carryover Functions ..... 46
3.3.1 The data ..... 46
3.3.2 A distributed lag model interpretation of the carryover functions ..... 47
3.4 Estimation of the Nutrient Non-substitution Model ..... 58
3.4.1 Alternative approaches ..... 58
3.4.2 Using linear splines to approximate crop response to individual nutrients ..... 61
3.5 A Test of the Nutrient Non-substitution Hypothesis. ..... 66
CHAPTER 4: EMPIRICAL RESULTS ..... 79
4.1 Fertilizer Carryover Functions ..... 79
4.1.1 The phosphorus carryover function ..... 79
4.1.2 The potassium carryover function. ..... 88
4.2 Crop Response Functions ..... 94
4.2.1 The polynomial formulations ..... 94
4.2.2 The nutrient non-substitution model ..... 105
4.3 Results and Discussion of Hypothesis Tests ..... 121
CHAPTER 5: CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH ..... 144
APPENDICES ..... 149
A. 1 Data Transformations ..... 149
A. 2 Application Rates for $P$ and $K$ on Purdue Fertility Experiment ..... 154
A. 3 Corn-1 Yield, Soil Tests and Applied Fertilizer ..... 155

Table of Contents (continued)
A. 4 Subroutine: CALCFG ..... 163
A. 5 Subroutine: CALCON ..... 163
A. 6 The MPS for the 1960 Non-substitution Problem ..... 164
A. 7 Solution for the 1960 Non-substitution Problem ..... 174
A. 8 Solution for the 1963 Non-substitution Problem. ..... 189
A. 9 Solution for the 1965 Non-substitution Problem ..... 198
BIBLIOGRAPHY ..... 207
3.1 Summary of the Field Operations at the Purdue Experiment ..... 48
4.1 Results for the Estimation of the Phosphorus Carryover - Function. ..... 82
4.2 Results for the Estimation of the Potassium Carryover Function ..... 89
4.3 Results for the Estimation of the Square Root Function. ..... 99
4.4 Results for the Estimation of the Quadratic Function. ..... 103
4.5 Estimated Spline Functions - 1960 ..... 114
4.6 Estimated Spline Functions - 1963 ..... 119
4.7 Estimated Spline Functions - 1965 ..... 120
4.8 C-Statistic Test Results ..... 125
4.9 Results for Auxiliary Estimations for the CP Testing Procedure - NS and SR ..... 128
4.10 Results for Auxiliary Estimations for the CP Testing Procedure - NS and Q. ..... 132
4.11 The Estimated CP Test Statistics ..... 133
4.12 Results for the Estimation of the Square Root Function- Pooled Data ..... 137
4.13 Estimated Spline Functions for the Pooled Data ..... 138
4.14 Hypothesis Test Results for the Two-Year Models: NS and SR. ..... 139
4.15 Statistical Status of Hypothesis Test Results Based on Different Testing Procedures ..... 141

## List of Figures

2.1 Depiction of the Law of the Minimum: the Case of Nitrogen and Phosphorus ..... 20
2.2 Depiction of the Mitscherlich Equation ..... 23
3.1 Approximating Crop Response to a Single Limiting Nutrient by Linear Splines. ..... 62
4.1 Extrapolation of a Single Season Proportionality Constant for Phosphorus ( $\lambda_{p}$ ) ..... 83
4.2 The Relationship Between the (Non-substitution model) Residual Sum of Squares $S(k)$ and the knots (k). ..... 112

## Chapter 1: Background

The importance of commercial fertilizers in agricultural production cannot-be over-emphasized. Farmers' enthusiasm about the success of fertilizers is clearly demonstrated by the fact that application rates for the three major nutrients - Nitrogen (N), Phosphorous (P), and Potassium (K) - have increased tremendously over the past three decades. In 1950/51, total world-wide use of $\mathrm{N}, \mathrm{P}$, and K was about 4.0, 2.6, and 3.3 million tons, respectively [FAO, 1952]. The corresponding consumption figures for 1978/79 are 51.0, 12.5, and 20.8 [FAO, 1979].

For the relatively immobile nutrients such as $P$ and $K$, recovery by crops during the growing season may be such that a sizable proportion of the nutrient is left over for succeeding crops. $1 /$ This left over portion of previously applied nutrients is commonly referred to as 'residual fertilizer' or synonymously as 'fertilizer carryover.' In subsequent chapters, it will be useful to refer to carryover in units of currently applied fertilizer; it will then be interpreted as the value of carryover.

Nutrient mobility, varying capabilities of crops to extract residual nutrients and repeated applications of the same fertilizer grades over the years have, in part, contributed to the accumulation of P and K in many agricultural soils.

Unlike $P$ and $K, N$ is relatively mobile. All forms of applied $N$ are usually transformed to nitrates $\left(\mathrm{NO}_{3}^{-}\right)$. Mobility of N is due, in part, to the transient nature of $\mathrm{NO}_{3}^{-}$. Nitrates are readily leached in moist and/or agriculturally disturbed ecosystems.

[^0]Nitrogen may also be lost by denitrification of $\mathrm{NO}_{3}^{-}$to nitrous oxide and nitrogen gases, particularly under anaerobic conditions of the soil. Global annual loss of N through nitrification and denitrification is estimated to be on the order of $50-60 \%$ of total soil supply [Wittwer, 1980]. Further losses of $N$ may be incurred by volatilization as ammonia gas.

It is not intended here to give the impression that all forms of $N$ are readily lost from all soil types. Nitrogen in the form of ammonium $\left(\mathrm{NH}_{4}^{-}\right)$is rather stable, and its conversion to $\mathrm{NO}_{3}^{-}$is dependent on soil and weather conditions [Likens et al., 1969]. Nitrogen carryover is, therefore, not uncommon.

The implication of the foregoing discussion is that nutrient carryover is a general problem which is by no means restricted to $P$ and $K$. It must, however, be admitted that, from an agricultural production point of view, the accumulation of $P$ is not only more prevalent, but is also more important than the case of the other major nutrients. The more mobile nutrients on the other hand pose a real danger with regard to environmental pollution.

Where carryover is significant, making recommendations to farmers purely on the basis of current applications of fertilizer not only leads to a misspecification of production costs but also to over-fertilization of the soil. The misspecification of costs is a consequence of the neglect of the contribution of nutrient carryover to current production. Over-fertilization, on the other hand, arises when further applications raise the total nutrient supply in the soil beyond the level required to achieve maximum crop response. It may also come about when other factors of production such as technology and management impose a ceiling on crop response. Excessive use of commercial fertilizers easily
translates into economic terms particularly in view of the rapidly escalating energy prices. It also leads to a waste of scarce mineral resources and unnecessary pollution of the environment.

One may ask the question: Why another research on crop response analysis and fertilizer use? The answer is simple: The era when extension agents solicited farmers to apply more and more fertilizers is over. This is now the time to answer these questions: Is there a response to added fertilizer? How do we avoid fertilizer wastage? Farmers often rely on fertilizer use recommendations issued to them by agronomists and agricultural economists. These scientists must therefore cooperate among themselves and with farmers if a meaningful solution to the fertilization problem is to be realized.

In the past, except for a few isolated cases, recommendations on fertilizer use have been derived from static optimization techniques. The basic physical relationship was the so-called production function relating crop yield and applied fertilizer. Fertilizer-output price ratios were then used to determine the optimal levels of application.

Increasing soil nutrient build-up, high and rapidly increasing energy prices, dwindling mineral reserves and the pressing issues of environmental pollution now suggest an urgent need for more comprehensive models. Such models must take into account past cropping patterns and yields, soil nutrient levels, soil types and weather variables in addition to the conventional input-output price ratios. A time dimension must be explicit in the models. Crop fertilization should, therefore, be viewed as a long-run dynamic problem. Such a conception should lead to efficient use of fertilizer while appropriately discounting and balancing future economic consequences. This obviously transcends the notion
of a production function as the basic tool for deriving fertilizer recommendations.

The motivations for the present study can now be summarized as follows: (a) the accumulation of the major nutrients, or excessive use of the same, in many prime agricultural soils and the need to popularize the specification and selection of comprehensive dynamic models to be employed in constructing optimal long term fertilization strategies and (b) the need for restructuring of research methodologies in order to obtain the requisite data for estimating such models.

The provision of optimal fertilizer recommendations to farmers generally entails three broad problem areas: (i) formulation and specification of the physical relationship between crop response and factors of production, (ii) collection and updating of experimental and/or survey data needed for estimation and selection of an appropriate specification and (iii) using the selected model for economic analysis. It is worth emphasizing that all three problem areas are interrelated and that they are equally important from an economic point of view.

This study will utilize experimental data in testing hypotheses regarding the formulation and specification of physical models of crop response to variable supplies of fertilizer. Economic analysis of fertilizer use under actual farming conditions will not be attempted. Once the physical models are verified, a localized application to an economic setup should be routine. The intention here is to concentrate only on the first and part of the second problem areas mentioned above. As will become apparent in subsequent chapters, the computational demands of these two problem areas are not slight.

In general terms, the objective of this study is to use statistical
techniques to test hypotheses regarding specification of the functional relationship between crop response and total soil supply of selected major nutrients. First, it is worth stressing that the intention is to employ statistical techniques as opposed to ad hoc and/or subjective criteria. Secondly, the study is concerned with testing of hypotheses rather than simple selection or discrimination among competing specifications. The latter approach usually leads to choice of one of the models as the best. Thus far, this has been the popular technique of selecting models for crop response analysis. Hypothesis testing, unlike model discrimination, need not always lead to a unanimous choice. It is conceivable that all models under test may be rejected in favor of an unspecified alternative. Finally, it should be noted that the variable of interest is total rather than applied nutrient.

The use of total nutrient supply is necessary whenever carryover is significant. A consistent estimate of total nutrient supply at a given time can only be constructed from consistent estimates of carryover. A secondary objective of this research is, therefore, to estimate the value of nutrient carryover. This can be done in a number of different ways, but for this study, cross-sectional and time series data on soil tests are pooled in a carryover function which provides consistent estimates of suitably defined conversion factors. The latter are then used to construct estimates of residual fertilizer.

A properly specified carryover function makes the task of estimating nutrient carryover and hence that of constructing total availability a lot easier. There is no need for terminating fertilizer applications in selected experimental plots in order to measure the rate of decline of a given nutrient. Furthermore, conversion factors obtained via the
carryover function, if assumed to be stable over a specified period of time, can be used in generating future estimates of total nutrient supply in a given soil-type under a particular cropping pattern. Additional historical information concerning the soil and cropping are then no longer required. All that is needed is the soil test which acts as a sensor of the stock of nutrient supply at the end of a growing season (i.e., prior to any fertilization).

The specification and estimating of crop response to total nutrients constitute another distinct sub-problem. The popular empirical approach has been to fit crop response data to a polynomial approximation by use of regression techniques. Exponential functions have also been used but, by and large, choice of the functional relationship has bordered closely on arbitrariness. In this study, the thesis is that a particular specification should be adopted only if it passes prescribed statistical tests. The set of specifications to be accorded serious consideration is, however, limited to those which conform to conventional theories regarding plant growth and the dynamic nature of the fertilization problem already alluded to.

Ordinary polynomials such as the square root and the quadratic are easy to work with and seem to fit empirical data well, particularly when yield is suddenly depressed beyond a certain level of fertilizer application. They are, however, unsuitable for examining relationships where the issue of substitution among the arguments is at stake. They are also inappropriate when the response surface is characterized by a significant plateau. Under such circumstances, an exponential function such as the Mitscherlich may be recommended. The general use of exponential functions in crop response analysis has, nevertheless, been criticized
on grounds of their inability to account for yield depression. Like any other continuous function, they do allow for varying degrees of substitutability among the macronutrients.

Choice of the mathematical form of the relationship between crop response and total nutrient supply has significant economic implications. For example, the concept of substitution has led some researchers to the conclusion that the same output can be produced by different combinations of a given level of resource endowments. The validity of such a conclusion when extended to major nutrients is questionable particularly in the face of significant nutrient build-up in many agricultural soils under continuous fertilization. The concept of substitution when applied to total nutrients may not be a useful one.

The rationale behind using total as opposed to applied nutrients as the basis for analysis has already been established. However, the question of whether or not to account for substitution and plateau-type crop surfaces is one which needs an empirical answer. It is hoped that this research will shed some light in that direction.

Lanzer [1978] employed a dynamic non-substitution model for economic analysis of fertilizer use in southern Brazil. Lanzer's model was a generalization of Liebig's 'law of the minimum.' The model was, however, chosen a priori, largely on the basis of biological considerations, despite the fact that the author pointed up the appropriateness of a statistical test of the non-substitution hypothesis. Truth of the non-substitution hypothesis is crucial to the validity of the mathematical specification of the model.

In this study, a dynamic non-substitution crop response model, similar to that employed by Lanzer, is to be estimated in a nonlinear
programming framework. Experimental data from the Agronomy Farm at Purdue University will be used. The programming approach to the estimation of a Liebig-type model may be more efficient in terms of data and computa-tional-requirements than the techniques used hitherto.

The final objective of the study is to carry out a test of the nonsubstitution hypothesis. The model under the alternative hypothesis will be a conventional continuous function, hence non-nested hypothesis testing procedures will be employed. Such testing procedures are appropriate whenever the hypotheses being tested do not belong to the same family, that is, when one cannot be obtained as a limiting case of the other.

The results of this study should have important implications, not only for experimental designs and data collection, but also for the selection of appropriate models to employ in the construction of long term fertilization strategies. Such strategies should lead to conservative rather than excessive use of commercial fertilizers in the future.

### 2.1 Formulation of crop response models

### 2.1.1 The need for a careful characterization of the response

## -surface

Traditionally, fertilization strategies have been derived from static optimization problems of the form

$$
\begin{equation*}
I=\operatorname{Max}_{X}\left[\operatorname{Pf}(X)-C^{\prime} X\right] \tag{2.1}
\end{equation*}
$$

The symbols are vector valued and the criterion function is some measure of profit, $\pi$. $P$ is the net value of the final product, $f(X)$, and $X$ represents the controllable inputs whose value is given by $C$.

The problem in (2.1) represents an idealized situation in which the influence of other uncontrollable factors have been isolated and the mean yield is a function of only the controllable variables. In its simplest form, $X$ is a vector representing factors such as the major fertilizer nutrients: Nitrogen (N), Phosphorous (P), and Potassium (K).

The above approach is typified by the works and writings of Heady and his associates at Iowa State College during the 1950's [see for example Heady et al., 1955]. Optimal levels of $X$ were derived from a relationship such as (2.1). In addition, isoclines, isoquants and other marginal relationships were computed to characterize the response surface.

Objections to the use of the optimization problem in (2.1) as a basis for designing recommendations to farmers have been numerous. The major ground for such objections is its oversimplification of nature. As used by the authors cited above and by many economists and agronomists thereafter, the objective function implies rather strong assumptions about
farmers' preferences. $\frac{1 /}{}$ It also neglects the dynamic and stochastic aspects of the fertilization problem.

One of the fundamental problems in crop response analysis, however, has been the specification of the functional form of $f(X)$ appearing in equation (2.1). The computation of the optimal level of $X$ is based on the derivative properties of $\pi$ and hence of the response function, $f(x)$. It is therefore hard to overemphasize the importance of a careful characterization of $f(X)$ as a basis for constructing economic recommendations to farmers. Different functional forms lead to different "optimum" levels of $X$. In a recent research in which the authors compared recommendation bias, mean square error and expected economic loss for the quadratic, square root, linear-and-plateau and Mitscherlich specifications using simulated data, Mombiela and Nelson concluded that "use of the wrong model would have serious adverse effects especially when considering regionwide application" [Mombiela and Nelson, 1980, p. 48].

A precise selection for the mathematical form of $f(X)$ must, however, depend on detailed experimental and historical evidence unique to a given set of data. Even when such evidence is available, the final choice can at best be a simplification of reality since crop yield is influenced by numerous factors and the "true" form of the relationship is unknown. Therefore, the use of the phrase "wrong model" as in the above citation would generally be inappropriate since it implies knowledge of the "true model." The question of model selection is a statistical one and is mentioned in its full right in Section 2.2. Model selection as discussed

[^1]in that section will be in reference to "a posteriori" judgments. The biological nature of plant growth can be exploited to screen off some of the specifications of $f(X)$ which do not conform to experience and certain fundamental considerations set up a priori. This task is taken up in the following sub-sections. It should, however, be noted that the separation of model selection (choice of a mathematical form) discussed in Section 2.2 from the formulation of crop response models (in the following sub-sections) is only for the sake of exposition; the two go hand in hand.

### 2.1.2 Polynomial approximations of crop response models

A generalized model of crop production may be stated as

$$
\begin{equation*}
Y=F(X, S, W, E, B) \tag{2.2}
\end{equation*}
$$

where $Y=$ crop yield
$X=$ vector of controllable variables
$S, W=$ vectors of soil-type and weather variables, respectively
$\varepsilon=$ vector of unspecified variables assumed to have a random influence on $Y$ and
$\beta=$ set of parameters of the system.
Most of the earlier efforts on crop response analysis have been preoccupied with the "production function" concept. Often the single equation analyzed took the form of

$$
\begin{equation*}
Y=f(X) \tag{2.3}
\end{equation*}
$$

where $Y$ is crop yield and $X$ is the set of controllable variables.
In isolated cases, soil-type and weather variables have been included explicitly in (2.3) as independent factors. Accounting for such variables, however, constitutes a distinct problem area which will be discussed in
sub-section 2.1.5. The present discussion will proceed under the assumption that the vector $X$ in equation (2.3) represents only such controllable factors like labor, planting density and fertilizer nutrients $N$, $P$, and K. But, it is not uncommon to find an economic analysis based on a response function utilizing only one variable. The need for including more variables in equation (2.3) arises not only from the fact that crop yield is a function of many variables but also from the fact that there may be interactions among such variables.

To be of use in economic analysis, the function $f(x)$ in equation (2.3) needs to be assigned an explicit mathematical expression. ${ }^{1 /}$ Since the "true" form of the function is unknown, it is possible to approximate it as a Taylor series expansion of $f(x)$. The expansion is done around some specified point $X_{0}$. If, taking a scalar simplification of $X, X$, $f(x)$ has a finite $n^{\text {t'n }}$ derivative $f_{n}(x)$ for all values of $x$ and $f_{n-1}(x)$ is continuous everywhere, Taylor's theorem states that for any specified value of $x, x_{0}$, and every $x=x_{0}$, a point $z$ exists in the interval joining $x$ and $x_{0}$ such that $f(x)$ may be approximated within a specified bound of error as

$$
\begin{align*}
f(x)=F\left(x_{0}\right) & +\sum_{k=1}^{n-1}\left[f_{k}\left(x_{0}\right)\left(x-x_{0}\right)^{k} / k!\right] \\
& +f_{n}(z)\left(x-x_{0}\right) / n! \tag{2.4}
\end{align*}
$$

A polynomial of the form

$$
\begin{equation*}
f(x)=\beta_{0}+\beta_{1} x+\beta_{2} x^{2}+\ldots \tag{2.5}
\end{equation*}
$$

1/This is not necessary when a discrete analysis as proposed by Hildreth [1954] is adopted. Choice of a mathematical form is only critical when the intention is to characterize the response surface by a continuous function and this has been the general tendency in crop response analysis.
is obtained by collecting like powers of $x$ in equation (2.4).
The use of polynomials in empirical economic analysis of crop response came in earnest in the 1950's. Heady and Pesek [1954] explored two polynomials (the quadratic and square root forms) and the Cobb-Douglas formulation.

Subsequent work on production functions particularly by Heady, Pesek, and Brown [1955] and Brown et al. [1956] were, methodologically, similar to the 1954 study cited above. The general conclusion was that the square root was an appropriate representation of crop response to fertilizer. Some of the selection criteria cited were the $R^{2}$, prediction of marginal products and prediction of isoquants and single variable-output curves. French [1956] compared five formulations and selected the square root and the Mitscherlich equations (the latter explained in the next subsection). The author concluded that choice between these two formulations would depend on computational convenience and thus, by implication, eliminated the Mitscherlich equation which belongs to the exponential family and hence entailed tedious computations compared to a polynomial.

The two volumes edited by Baum, Heady, and Blackmore [1956] and Baum et al. [1957] summarize the cooperative efforts among agricultural economists, agronomists, and statisticians. Such efforts were aimed at specification and estimation of production functions and the design of recommendations leading to efficient use of fertilizer. The credibility and influence of these two volumes and the myriad of ensuing publications in the 1950's and 1960's was overwhelming. Thus, economic analysis of fertilizer use in the last three decades has been dominated by polynomial approximations of one form or the other.

Within the family of polynomials, other researchers have preferred
the quadratic to the square root even when the latter gave a better "fit" to the data. Tollini and Seagraves [1970], for example, employed the quadratic equation despite evidence pointing in favor of the square root. Their reasoning was that the linear term in the square root is expected to be negative but is empirically found to be close to zero or positive. The authors also stated that the linear term in the square root equation is more complex or difficult to explain particularly in large equations.

Despite the cooperative efforts encouragingly expressed by Glenn Johnson [in Baum et al., 1957], there is still a lack of interdisciplinary consensus of opinion regarding the choice of a mathematical form to describe crop yield data. Exponential functions, while being penalized by many researchers on account of computational inconvenience, have all the same found their way into the literature through the patronage especially of agronomists and soil scientists. Typical examples are Bray [1954], Hanway and Dumenil [1955], Engelstad and Terman [1966], and Helyer and Godden [1977]. The immediate question that arises is: If polynomials "fit" yield data well and are also simpler to estimate, why do some researchers continue using more complicated mathematical forms?

The question posed above is difficult to answer without resorting to historical and biological evidence pertinent to a given set of data. Such evidence is helpful in rationalizing the propositions and assumptions which constitute competing models of crop response. A polynomial approximation, with its underlying theories and assumptions, is one such model. Different theories and assumptions will lead to different models some of which may be more or less mathematically complex than a polynomial. It should be emphasized here that the statistical problem of choosing among competing models usually concerns testing of the validity of the
underlying theories and assumptions. The problem of model selection on the basis of statistical criteria is deferred to Section 2.2. In the next sub-section, non-statistical evidence is given which indicate that a polynomial approximation of crop response has serious shortcomings. Its simplicity, which was an asset before electronic computers became common-place, may no longer be a desirable attribute.

### 2.1.3 Objections to polynomial approximations

The two polynomial approximations popularly used in crop response studies have been the quadratic and the square root forms. Both exhibit diminishing and negative return phases. Although the square root has a slightly flatter surface of the $X$-space, they both display unique maxima and symmetry of curvature around such maxima. The ratio of the maginal productivities of $X_{1}$ and $X_{2}$ (in a two input case), commonly referred to as the rate of technical substitution, declines at a constant rate for the quadratic while for the square root it declines at a decreasing rate. The isoclines of both forms converge at a point in the $X$-space but those of the quadratic are not forced through the origin and are linear unlike in the case of the square root.

Some of the properties mentioned above are desirable from the point of view of what experience in crop production indicates. However, polynomial approximations have certain undesirable mathematical properties. Putter et al. [1966] recommended the use of equations containing "meaningful" parameters instead of polynomials whose parameters are not readily interpretable. Nelder [1966], in suggesting use of inverse polynomials, criticized ordinary polynomials for unboundedness which makes the task
of extrapolation difficult. The author added that the arbitrary nature of polynomials, justified only on grounds of maximization of $R^{2}$, precludes generalization beyond the specific set of data used for parametric estimation. Terman and Nelson [1976] also stated that the rather confused state of plant analysis interpretation may be attributable to the use of quadratic regressions since the numerical values of the coefficients do not correspond to biologically meaningful entities. In a more recent paper, White [1980] has criticized the Taylor series approximation interpretation of ordinary least squares (OLS). The author stated that tests of hypotheses may be seriously misleading and that OLS estimates may not necessarily provide reliable information about local (derivative) properties of unknown functions, except under very restrictive conditions.

On a more general perspective, functional specification in crop response analysis (but not necessarily choice among competing forms) has been guided largely by biological principles of plant growth. Some of the fundamental principles which are normally considered are: (i) nutrient essentiality and issues relating to interaction and substitution among macronutrients, (ii) the exponential aspect of biological growth, (iii) diminishing increments in production; the basis of which conforms to the concept of diminishing returns in economics, and (iv) the concept of nutrient mobility and the residual value of nutrients.

In agricultural production, Nitrogen (N), Phosphorous (P), and Potassium (K) are the most important essential nutrients. Essential nutrients do not substitute for one another physiologically. This fact is widely acknowledged even by strong proponents of polynomials. There is, however, the concept of economic substitution justified for a variety of reasons. Pesek and Heady [1956, p. 243] for instance stated that:

It is known, of course, that substitution of nutrients such as $\mathrm{P}_{2} \mathrm{O}_{5}$ and $\mathrm{K}_{2} \mathrm{O}$ does not take place in the chemical process of the plant. However, smaller or larger quantities of one nutrient affect the availability of the other nutrient in the soil or alter the composition of the plant. In this sense, several combinations of nutrients allow the same yield.

This view seems to have been supported by many others. Munson and Doll [1959, p. 153], for example, stated that "This [substitution] is probably due to the effect that the addition of one nutrient can have on the utilization of other soil nutrients." It is therefore apparent that the concept of economic substitution is derived from nutrient interaction and/or is applicable only to cases where the analysis is based on applied rather than total (available) nutrients. When applied to the former condition, the idea is misleading (as explained shortly) and is short-sighted in the latter case. That is, ignoring soil available nutrients and then talking of economic substitution among applied nutrients is not a sound long-run proposition.

Interaction among variables affecting plant growth and indeed among the essential nutrients, $N, P$, and $K$ is an indisputable fact. This does not mean that interaction is always there. The existence of interaction between two variables is an empirical matter. When polynomials are used to analyze yield data, the inclusion of cross-product terms such as $X_{1} X_{2}$ in the equation is often used as a means of investigating the presence of such interaction effects. Interaction, as explained below, is distinct from substitution. But, it should be noted that substitution is an intrinsic mathematical property of polynomials and that it is inappropriate to use such functional forms to investigate the existence of substitution among variables.

Let $X_{i j}$ represent the $j^{\text {th }}$ level of the $i^{\text {th }}$ nutrient and $f\left(X_{i j} \mid X_{i}{ }_{i j}\right)$
the crop yield obtained from application of $j$ units of the $i^{\text {th }}$ nutrient given $j^{\prime}$ units of the $i^{\prime}$ th nutrient. Empirically, it is said that there is no interaction if $f\left(x_{11} \mid x_{20}\right)-f\left(x_{10} \mid x_{20}\right)=f\left(X_{11} \mid x_{21}\right)-f\left(x_{10} \mid x_{21}\right)$ or alternatively, when $\Delta f\left(X_{10} \mid X_{20}\right)=\Delta f\left(X_{11} \mid x_{21}\right)$; otherwise there is interaction between $X_{i j}$ and $X_{i ' j}{ }^{\prime}$. Interaction can be negative or positive and as can be deduced from the alternative formula given above, implies that, in the limit, $\partial f(X) \mid \partial X_{i}$ will be a function of $X_{i}$..

A polynomial formulation would, of course, display interaction but not without introducing substitution. One of the functional formulations which does not permit substitution is based on the "law of the minimum" formally stated by Liebig in 1840.

The basis of the law of the minimum is nutrient essentiality. As expounded by Liebig, "the crops on a field diminish or increase in exact proportion to the diminution or increase of the mineral substances conveyed to it in manure" and that "by the deficiency or absence of one necessary constituent, all others being present, the soil is rendered barren for all those crops to the life of which that one constituent is indispensable" [Russel, 1950, p. 13]. The latter quotation has not been contested. The former has been disputed since it suggests an unbounded function of the form $f\left(X_{i}\right)=\beta_{0}+\beta_{1} X_{i}$ for given constants $\beta_{0}$ and $\beta_{1}$ :

There have been various modifications of Liebig's law. Some of the modifications have resulted in exponential functions (to be discussed shortly) but there are analysts who contend that "Liebig's original law of the minimum still stands as a logical and simple theory, meriting a statistical method for fitting it to observations" [Waggoner and Norvell, 1979]. The authors combined the two aspects of the law relating to proportionality and deficiency to obtain the function

$$
\begin{equation*}
y=\operatorname{Min}\left[f\left(z_{1}\right), f\left(z_{2}\right), m\right] \tag{2.6}
\end{equation*}
$$

where $y=y i e l d$

$$
\begin{aligned}
f\left(z_{i}\right) & =c_{i}+d_{i} z_{i} ; i=1,2 \\
m & =\text { maximum yield obtained by adding the nutrients } z_{i} \\
c_{i}, d_{i} & =\text { constants } .
\end{aligned}
$$

The function says that yield is limited by either $Z_{i}$ or $m$ or some combination thereof, but $y$ is only a function of the limiting factor which is the meaning of the expression $\operatorname{Min}(\cdots)$. Thus, if $Z_{1}$ and $Z_{2}$ are not limiting growth, $y=m$. Figure 2.1 depicts the law of the minimum.

The law of the minimum embodies not only a function that allows interaction without substitution but also the idea of a yield plateau. The existence of a yield plateau may be attributed to a number of reasons some of which are:
(a) The elements $N, P$, and $K$ do not cause yield depression for most crops unless the application rate is extended substantially beyond the point of maximum response
(b) A ceiling on crop yield may be posed by other factors such as management, technology and unspecified factors of production
(c) The exponential nature of most biological responses.

Tollini and Seagraves [1970] stated that "no one knows the form of the true production function but experience and theory would suggest that it has a flatter top and a sharper bend than is possible with the quadratic."

The existence of yield plateaus have also been documented empirically. Boyd [1970] reported that plateaus were reached at low or intermediate

Figure 2.1 Depiction of the Law of the Minimum:

rates of fertilizer for a number of different crop-nutrient combinations. Anderson and Nelson [1975] have also reported that most of the Iowa State College data, used extensively in the 1950's to popularize polynomial approximations, exhibit plateau-type maxima. The authors suggested formulations or analytical techniques such as the use of splines because the quadratic polynomials overestimate both fertilizer level and yield corresponding to the highest net profit.

The intention here is not to give the false impression that all fertilizer-crop combinations exhibit yield plateaus. The point to be emphasized is that in cases where plateaus do occur (and theory and experience attest to a preponderance of such cases), ordinary polynomials lead to positive biases in fertilizer recommendations. This is a fact that cannot be taken lightly particularly in environments where pollution from agricultural production is a pressing issue. It is, however, of universal importance in view of the rising cost of energy and the need to conserve scarce mineral resources that go into the production of commercial fertilizer.

The foregoing discussion suggests that what is needed is a functional form that will account for the biological phenomena such as interaction and plateaus in crop response surfaces. The law of the minimum is a step in the right direction. But the aspect of it claiming that crop yields diminish or increase in exact proportion to the decrease or increase in available plant nutrients is contrary to the "theory of diminishing increments." This theory is based on the fact that yield from fertilizer usually increases at a decreasing rate. A formal statement of the theory is attributed to Mitscherlich [1909] and Spillman [1923]. Mitscherlich's version of the theory (referred to hence as the Mitscherlich equation)
is based on the assumption that there exists a yield maximum, $A$, and that yield increments are proportional to the decrement from the maximum. The equation, for $k$ nutrients, is stated as

$$
\begin{align*}
Y=A(1 & \left.-e^{-c_{1} X_{1}}\right)\left(1-e^{-c_{2} X_{2}}\right) \cdots\left(1-e^{-c_{k} X_{k}}\right)  \tag{2.7}\\
\text { where } Y & =\text { yield } \\
A & =\text { yield maximum } \\
X_{i} & =\text { total supply of deficient nutrient, } i=1, k \\
\text { and } c_{i} & =\text { fertilizer efficiency constant for the } i^{\text {th }} \text { nutrient. }
\end{align*}
$$

The Mitscherlich equation is a limited substitution model. This implies that in an $X_{1}-X_{2}$ plane, the isoquants have a restricted region over which the elasticity of substitution of one variable for the other is different from either infinity ( $\infty$ ) or zero. This is in contrast to polynomials which usually portray smooth isoquants with a wide range of substitutability.

For a single nutrient, a depiction of the Mitscherlich equation is given in Figure 2.2 which also shows a curve for a second degree polynomial such as the quadratic. The illustration also shows a curvilinear-and-plateau (CLP) model which is similar to the Liebig function except that the former is restricted to account for the law of diminishing increments.

Assuming a plateau-type crop response surface, the nature of misspecification which would arise from the use of a polynomial should be clearly apparent from Figure 2.2. On the other hand, if yield were depressed after a certain level of fertilizer application, the other formulations shown in Figure 2.2 would be unable to capture it. However, as already noted, for essential nutrients, such yield reductions usually

Figure 2.2 Depiction of the Mitscherlich Equation

Yield

occur at levels of fertilizer application well beyond the feasible economic range. Hence, the use of the Mitscherlich or the CLP model when yield depressions occur may not constitute problems as serious as the use of a polynomial in cases where the surface has an extensive plateau. It must be emphasized that all the three models constitute a potential bias of one kind or the other. Choice among them may therefore be on the basis of whichever does the least damage to theory and experience in crop production.

Bray [1954] in introducing the nutrient mobility concept suggested that mobile nutrients may follow Liebig's law while the relatively immobile nutrients may act according to the Mitscherlich-Baule percentage sufficiency concept. Swanson [1963] employed Bray's concept of nutrient mobility to combine nonlinearity in response (implicit in the Mitscherlich formulation) with Liebig's law of the limit. The author explained how to go about obtaining a linear programming solution to a problem in which the response to one of the two nutrients being studied was nonlinear. The author also noted that to estimate the yield response to one nutrient, it would be necessary to make sure that the unstudied factors are not effectively limiting production. Whether this assurance is to be achieved by experimental design or by analytically discriminating among the observations was not made clear. Empirical work along Swanson's footsteps has been limited.

Anderson and Nelson [1975] suggested a family of intersecting straight lines (linear splines) to approximate the feasible economic range of a crop response curve. In Chapter 3 it will be shown how the idea of splines can be utilized in conjunction with Liebig's law of the limit to obtain a generalized non-substitution crop response model.

### 2.1.4 Incorporating site and weather characteristics in crop response functions

The incorporation of site and weather variables in comprehensive crop response models has been as controversial as it has been slowed by the dearth of suitable data. The importance of general crop response functions accounting for crop variety, soil and weather variables is well documented in literature (see e.g., Hildreth, 1957).

The concept of nutrient carryover and the increasingly menacing problem of pollution arising from the use of commercial fertilizers have motivated research in the direction of exploring the dynamic nature of crop-fertilizer relations. Lack of recognition of the dynamic nature of the fertilization problem can lead to costly long-run problems.

One way of accounting for residual fertilizer (carryover) is through soil tests. The latter, however, do not translate one-to-one into equivalent amounts of added fertilizer. Their inclusion in crop response functions has therefore been handled in various ways. One general approach is to represent the mean response by $f\left(X^{\top}\right)$ where $X^{\top}$ is the total nutrient given by $\lambda b^{*}+X$, where $b^{*}$ is the soil test, $\lambda$ is an unknown constant of proportionality and $X$ is the applied nutrient. This approach was employed by Hildreth [1957]. Soil tests may, however, be highly interdependent in addition to being functions of crop yield and added fertilizer nutrients. Hence, a generalization of Hildreth's approach as carried out for instance by Jensen and Pesek [1959 a \& b] and Ryan [1972] can lead to statistical and interpretational problems.

Another problem encountered in the use of soil tests is that they may not be correlated with the plant extractable amounts of nutrients under all circumstances. It may therefore be appropriate to look into
implicit ways of measuring the residual value of past applications of fertilizer.

The idea of including soil tests explicitly into the crop response model is attractive because such tests can be readily conducted to monitor the fertility level of a soil at a given time without requiring a record of past fertilizer applications. The latter are indispensable if carryover is computed explicitly.

The soil-type and weather variables problem has also been tackled in alternative ways. The difficulties usually encountered are the selection of variables to include in the regression, the accuracy of measuring such variables and the interpretation of the results. Steinberger [1957] for example, used dummy variables in a discrete model suggested by Hildreth in 1957 while Ryan [1972] included soil pH, temperature, and rainfall in addition to fertilizer nutrients in a single quadratic equation. Models of the latter type run the risk of muticollinearity as already noted above in relation to soil tests.

Lanzer [1978] employed a generalized crop response model requiring no explicit assessment of soil-type and weather variables. The virtues of his approach derive from the fact that these two categories represent a large number of variables which would be costly if measured on a largescale basis and statistically worthless unless collected over a long period of time.

### 2.2 Testing Model Specification

The importance of choosing an appropriate functional form for the crop response function, $f(X)$, has already been stated in subsection 2.1.1.

But, there is no basic theory to indicate what this functional form should be. It is because of this reason that it was stated earlier that the precise selection of any mathematical function to express $f(X)$ depends on detailed experimental evidence.

The phrase 'experimental evidence' should be taken to encompass the aspects discussed in sub-sections 2.1.2 through 2.1.4. It refers to the body of prior information (mathematical, historical, or biological) that helps the analyst in delineating the possible configurations of the response surface. Such evidence cannot be used unilaterally since the principles upon which they are based are not unswerving universal laws. Indeed a selection procedure based entirely on the analyst's evaluation of the underlying circumstances may entail a significant degree of subjectivity. Theil [1971, p. 545] states that "statistical procedures should not be regarded as the only tools for handling the selection problem. The analyst may be convinced on a priori grounds that one specification is more realistic than another, in which case he should feel justified in applying the former . . . ." It follows then that prior information and statistical criteria play a major role in model selection. The two go hand-in-hand. That is why some effort was diverted to an exposition of the problems akin to formulation of crop response models (Section 2.1) even though a statistical test of functional specification is one of the major goals of this study.

The term "statistical" as used here should be taken to mean any post-estimation criteria that might be used in model selection. They are criteria derived from the data according to some specified mathematical formulae. Thus economic procedures of selection are viewed as statistical.

It is common practice to compare different functional forms on the basis of the coefficient of multiple correlation ( $R^{2}$ ) and/or the estimate of residual variance, both properly adjusted for degrees of freedom. The function that maximizes $R^{2}$ or, alternatively, the one with the minimum residual variance estimate is taken as the "best" [Theil, 1971]. Other such criteria which aim at obtaining the single most appropriate formulation are: nesting of functions (preferably linear) into a comprehensive model and applying significance tests such as t-ratios on specified coefficients; maximum likelihood tests and cross validation. In the latter procedure, the data set is split into two segments (usually arbitrarily), one for model estimation and the other for validation.

The ultimate choice of a model cannot be based purely on the basis of the general ability of the model to represent the data even if this means representation in the vicinity of the optimum. The reason is that choice on the basis of ad hoc statistical criteria mentioned above can lead to deadlocks particularly in cases where most of the models being tested account for a high proportion of the variation in the dependent variable. Tollini and Seagraves [1970, p. 9] for instance, had a case in which all the three models being tested had $\mathrm{R}^{2}$ values of about $82 \%$ while Lanzer [1978] presented results of a comparison in which the square root and the Liebig-Mitscherlich models were indistinguishable solely on the basis of $R^{2}$. In the latter case, additional information such as residual variance estimate or more subjective criteria such as simplicity of the model and computational convenience would have to be considered in the decision process. Such cases are not uncommon in practice.

The use of ad hoc statistical procedures lead not only to practical difficulties but also to conceptual problems. Generally, the procedures
lead to a single formulation as the "best." For example, in the use of residual variance estimate ( $s^{2}$ ), it is assumed that the correct specification $f\left(X_{0}\right)$ has the minimum residual variance, $S_{0}^{2}$, and hence that any other specification $f\left(X_{1}\right)$ has $S_{1}^{2}$ such that $E\left(S_{1}^{2}-S_{0}\right)>0[$ Theil, 1971, p. 543]. $\frac{1 /}{}$ Selection on the basis of the specification with the minimum $s^{2}$ presumes that the form being tested is the true specification. Thus, in testing a maintained hypothesis ( $H_{0}$ ) against an alternative one $\left(H_{1}\right)$, the procedures imply conviction about the truth of $H_{0}$. In practice, there are not many cases in which the analyst would be certain about the truth of a given formulation. As already repeated many times, the true nature of the crop response function $f(X)$ is unknown. Hence, maintaining one hypothesis as the unequivocal truth or, alternatively, a selection procedure that seeks just one model as the "best" is inappropriate.

It is often stated that even when models are indistinguishable on the basis of such statistical criteria as $R^{2}$, they may still differ in terms of prediction and derivable economic recommendations. Some analysts have, therefore, preferred to conduct the model selection process on the basis of economic yardsticks. An example is when models are discriminated on the basis of a measure of economic return such as profit associated with fertilizer recommendations from the models. Perrin [1976] employed a similar procedure to compare linear response and plateau (LRP) and quadratic formulations. The author found no significant difference between the two formulations on the basis of economic grounds. The choice between the two would, therefore, have to be assisted by additional, possibly subjective, criteria. This is the kind of deadlock that is often

[^2]encountered when attempting to choose a single "best" formulation on the basis of "a posteriori" statistical formulae. The existence of this problem is, however, a blessing in disguise since it helps to underscore the importance of the fact that statistical criteria alone are not sufficient as a basis for model selection.

An extension of Perrin's method has been done for example by Mombiela and Nelson [1980]. The authors compared quadratic, square root, and Mitscherlich formulations on the basis of closeness to the true optimal recommendation. The true model was developed by the authors. Sampling from the model was then carried out to create data sets that were used to estimate the equations to be compared. The authors concluded that the Mitscherlich had the best performance and that "The quadratic surface is too rigid to accommodate non-symmetrical shapes . . . ."

The above procedure depends critically on the existence of an accumulated amount of experience gained through years of fertilizer research and practice. Without such experience (and this would most likely be the case in many developing countries), it would be impossible to construct the so-called true model from which the data is generated. Therefore, this technique is of limited applicability.

A general criticism of economic criteria is that their use in comparing models, one of which is misspecified, is not valid. The reason is that biases in the physical model are manifested in the economic testing procedure and thus minimizing the power of economic criteria as a means of testing model specification. A typical example is the comparison of an exponential formulation such as the Mitscherlich with a quadratic polynomial known to overestimate both the optimal level of fertilizer application and the corresponding crop yield when the response is
characterized by a plateau-type surface. In such a case, selection on the basis of economic grounds or the use of statistical techniques involving conventional likelihood ratio and $F$ tests are inappropriate. This is due to the fact that the comparison involves non-nested hypotheses. Two hypotheses are non-nested when the specification of one cannot be obtained from the other by imposing appropriate restrictions or be obtained as a limiting form of a suitable approximation [Fisher and McAleer, 1981].

The selection criteria reviewed above involve testing a maintained (null) hypothesis which is either rejected or not rejected. It is presumed that one of the hypotheses is true, that is, it obeys all the classical assumptions. First, because the true nature of the crop response function is unknown, it is difficult to develop a prior conviction about the truth of a particular hypothesis. Secondly, in crop response studies, many of the comparisons involve models like polynomials and exponential equations which do not belong to the same family. These points imply that the model selection procedure should be flexible enough to allow the possibility of rejecting both the maintained and the alternative hypotheses. This means that model selection should be preoccupied with hypothesis testing not merely discrimination in which the "best" model is sought. Since the models often do not belong to the same family, the hypothesis testing cannot proceed on the basis of measures of relative fit such as $R^{2}$ and residual variance estimate.

In this study, testing model specification is to be carried out by means of statistical procedures which take into account the non-nested nature of the hypotheses under test. Alternative models of crop response and the theoretical concepts of testing the underlying hypotheses are
presented in the next chapter. In the chapter dealing with empirical results (Chapter 4), the main focus will be on non-nested hypothesis tests and the facilitating procedures. Whenever possible, the present habit in the'econometrics literature of reporting relative measures of performance (e.g., $R^{2}$ and residual variance estimate) of the models studied will be sustained. However, to avoid distracting attention from the main objective of this study (namely, to carry out a test of the nutrient nonsubstitution hypothesis) and to avoid unnecessary duplication of previous research efforts, discussion of the results of estimation of the competing models will be suppressed to a minimum.

### 3.1 The nutrient non-substitution crop response model

The problems of variable selection and testing model specification by use of prior and statistical information have been considered in Section 2.2. Objections to the use of polynomials to represent the crop response function, $Y=F(X, S, W, E, \beta)$, given in equation (2.2) were discussed in subsection 2.1.3. It was also argued in subsection 2.1.4 that it is more appropriate to consider the fertilization problem as dynamic in nature. This requires an explicit account of the interrelationships between past and current events and their impact on crop response. Of special interest are the uncontrollable soil (S) and weather $(W)$ variables and efficient techniques of incorporating them into the models used for designing recommendations on optimal levels of the controllable factors $X$.

There are numerous soil and weather variables affecting crop response. Hence, an attempt to expand the number of relevant explanatory variables included in the functional relationship soon runs into statistical and measurement problems. To avoid the necessity of explicitly accounting for the multitude of soil-type and weather variables, Lanzer [1978] employed an alternative approach. $\frac{1 /}{}$ If $X$ in the general crop response function is taken to represent only fertilizer nutrients, the assumption of weak separability (between soil-type and weather variables on the one hand and fertilizer nutrients on the other) allows equation (2.2) to be written in the form

$$
\begin{equation*}
Y=g(S, W) f(b+X) \tag{3.1}
\end{equation*}
$$

[^3]In equation (3.1), b represents a vector of nutrients initially available in the soil and $g(\cdot)$ and $f(\cdot)$ are functions of variables whose symbols have already been explained. For simplicity of presentation, the error term ( $\bar{\varepsilon}$ ) is omitted from (3.1).

The response surface is assumed to be characterized by a yield plateau representing the maximum yield, $A^{*}$. The latter is dependent on a set of ideal conditions $(b+X) *, S^{*}$ and $W^{*}$ such that

$$
\begin{equation*}
g\left(S^{*}, W^{*}\right) f(b+X)^{*}=A^{*} \geqq g(S, W) f(b+X) \tag{3.2}
\end{equation*}
$$

by dividing both sides of (3.1) by $A^{*}$ one obtains

$$
\begin{align*}
& Y=A^{*} h(S, W) f(b+X)  \tag{3.3}\\
& \text { where } h(S, W)=g(S, W) / A^{*}
\end{align*}
$$

Since by (3.2) $A^{\star} \geqq Y$, equation (3.3) implies that $h(S, W) \in[0,1]$ and $f(b+x) \in[0,1]$. Thus, (3.3) is a relative yield model. By letting $A_{\text {sw }}$ represent $A^{*} h(S, W)$, equation (3.3) is written as

$$
\begin{equation*}
Y=A_{S W} f(b+X) . \tag{3.4}
\end{equation*}
$$

The term $A_{S W}$ is the yield plateau of a given set of experimental data as determined by soil-type and weather conditions at the site. One of the advantages of the formulation in (3.4) is that the latter variables need no longer be measured explicitly. The maximum yield for a given experiment may be used to code the data leaving only the problem of specifying the mathematical form of $f(b+X)$.

As already noted, the major building blocks $N, P$, and $K$ do not substitute for each other in a physiological sense. Letting $X$ represent the essential nutrients named above and invoking Liebig's law of the minimum, the relative yield model can be written as

$$
\begin{equation*}
Y=A_{S W} \operatorname{Min}_{i \in I}\left\{f_{i}\left(X_{i}^{\top}\right)\right\}+\varepsilon \tag{3.5}
\end{equation*}
$$

```
where }Y=\mathrm{ crop yield
    x
    I = the set of essential nutrients
    A
and \varepsilon = an error term.
```

Equation (3.5) is the nutrient non-substitution model. It differs from (3.4), which is merely a generalization of the Mitscherlich equation, in the sense that whereas the latter (and hence 3.4) allows limited substitutability, (3.5) does not. The model generalizes the one used by Waggoner and Norvell [1979] in the sense that the functions, $f_{j}\left(X_{j}^{\top}\right)$ are not restricted to be linear. According to the theory of diminishing increments, one would presume that these functions are concave. This is a useful assumption particularly when it comes to estimating the nonsubstitution model as will be shown in Section 3.4.

### 3.2 Formulation of the nutrient carryover functions

The units of $b_{j}$ as introduced in equation (3.5) need further explanation. The variable $b_{i}$ is interpreted as the amount of the $i^{\text {th }}$ nutrient already in the soil. It is the flow component of the stock of the $i^{\text {th }}$ nutrient carried forward from past growing seasons and is invariably referred to in the literature as nutrient carryover or the residual value of past fertilizer applications. The units are usually in kg of freshly applied fertilizer to which the carryover is currently equivalent.

The notion of $b_{i}$ as the amount of seasonal nutrient release suggests that it may be a function of factors such as the soil's capacity to hold the nutrient, the kind of crop grown, the duration of plant growth, the
prevailing weather conditions and the initial fertility level, to mention but a few. Nutrient carryover is therefore best viewed as stochastic. This may explain, in part, the seasonal variability in crop response to applied fertilizer as documented by agronomists [see for example Matocha et al., 1970].

In Section 3.1, crop yield was presented as a function of total amounts of soil available nutrients, $x_{j}^{\top}$. The latter is the sum of the carryover $\mathrm{b}_{\mathrm{i}}$ and freshly applied fertilizer $\mathrm{X}_{\mathrm{i}}$. Knowledge of the amount representing $b_{i}$ is therefore crucial to economic optimization of fertilizer applications over time. Unfortunately, no scientific means are available yet that can measure the variable directly. A chemical analysis of the soil (soil test) at the beginning of the season can, however, give an indication of the potential amount of a nutrient, part of which may be made available to the current crop. A soil test does not give the nutrient release or carryover as explained above. It merely gives an index which is, hopefully, correlated with what the crop can extract from the soil to augment the freshly added fertilizer. It follows then that the soil test variable, call it $b_{i}^{*}$, is also dependent on the soil-type, agronomic and weather variables mentioned in the preceding paragraph.

The dependence of $b_{j}$ and hence $b_{i}^{*}$ on soil-type variables is what makes soil testing valuable. On the basis of calibrated soil tests, soils are categorized according to their nutrient releasing ability or equivalently by use of an index reflecting $b_{i}$. This in turn enables researchers to localize fertilizer recommendations to specific soil types. The economic justification of such an exercise, however, hinges on the existence of a scientific procedure that produces $\mathrm{b}_{i}^{*}$ values which best reflect that component of the $i^{\text {th }}$ nutrient which is extractable by the
crop. The cost of fine-tuning fertilizer recommendations on the basis of soil tests must also be weighed against that of a blanket-type approach based on the average of all soil test results [see for example Ryan, 1972, ch. 5]:

As already noted above, soil tests give an index $b_{i}^{*}$ of the possible nutrient release $b_{i}$ in the current season. In order to utilize the soil test index, it must be converted to the more useful variable $b_{i}$. The rest of this section discusses the relationship between $b_{i}$ and the soil test index and how the current nutrient release is related to the past events: the concept of a nutrient carryover function.

Generally, the function which has to be specified is of the form $\mathrm{b}_{\mathfrak{i}}=\mathrm{h}\left(\mathrm{b}_{\dot{j}}^{*}\right)$. Reuss and Geist [1970] proposed the use of

$$
\begin{equation*}
h\left(b_{i}^{*}\right)=\lambda_{i} b_{i}^{*} \tag{3.6}
\end{equation*}
$$

In (3.6), $\lambda_{i}$ is the (unknown) proportion of soil available nutrient released in the current season. It is the proportion of the stock that will be extracted by the incoming crop; the stock is given by the index $\mathrm{b}_{\mathrm{i}}^{*}$. The authors went on to suggest that the function $\mathrm{h}\left(\mathrm{b}_{\dot{i}}^{*}\right)$ may be represented by more complex specifications such as a second degree polynomial in $b_{i}^{\star}$.

Mombiela and Nelson have cited a study by Mombiela et al., [1980] in which the authors found a linear relationship of the form $h\left(b_{i}^{*}\right)=a_{0}+a_{1} b_{i}^{*}$ to be more appropriate [Mombiela and Nelson, 1980, p. 2].

By considering the nutrient release as coming from distinct sources, e.g., a depletable source and a non-depletable source (the contribution of the native soil material), it is possible to end up with more complex functional forms for $h\left(b_{i}^{*}\right)$. Heylar and Godden [1977] have employed one such formulation.

Lanzer [1978] in a study based on data from a wheat-soybeans rotation in southern Brazil, used the simpler version given in equation (3.6) above. This simpler relationship will be adopted here, admittedly to facilitate ease of estimation of subsequent models relying on parameter estimates derived from the function $h\left(b_{i}^{*}\right)$.

For a given season and crop, the nutrient release depends on soil type [see for example Fixen and Carson, 1978]. One would therefore expect the $\lambda$-values to be a function of the soil characteristics. Equation (3.6) is therefore modified to

$$
\begin{equation*}
b_{i}=\lambda_{i s} b_{i}^{\star} \tag{3.7}
\end{equation*}
$$

The soil test ( $b_{i}^{*}$ ) is given in $\mu \mathrm{g} / \mathrm{g}$ while the nutrient carryover ( $\mathrm{b}_{\mathrm{j}}$ ) is in kg of applied fertilizer to which the nutrient release is equivalent. The subscripts $i$ and $s$ refer to nutrient and soil-type category, respectively.

Assuming that the $\lambda$-values are stable from year to year, soils may be categorized on the basis of their $\lambda$-values. The optimal level of fertility required to sustain adequate crop growth in a specified soil-type can then be computed only on the basis of current soil test and the corresponding estimate of the unknown proportionality constant, $\lambda_{\text {is }}$. $A$ record of past fertilizer applications at a particular location is therefore not required in designing future fertilizer recommendations to farmers. This is the sole purpose of soil tests: to be able to calibrate and use them for monitoring the soil fertility. ${ }^{1 /}$
$\underline{1 /}$ Since $y=f\left(X_{i}^{\top} ; \beta_{j}\right)=f\left(b_{i}+X_{i} ; \beta_{i}\right)$ and $b=\lambda_{i s} b_{i}^{\star}$ it follows that soil tests can also be calibrated on the basis of the efficiency of carryover fertilizer given by $k_{i}=\beta_{i} \lambda_{i s}$ where $\beta_{i}$ is the efficiency

The phrase "fertilizer recommendations" has been mentioned in the foregoing discussion but only for the sake of facilitating an understanding of the role of $\lambda_{\text {is }}$. Fertilizer recommendation is in the domain of economic optimization techniques and transcends the estimation of the underlying physical relationships as intended here.

Attention will now be directed towards a technique for obtaining an estimate of $\lambda_{\text {is }}$. It should be remembered that $\lambda_{\text {is }}$ simply facilitates transformation of the soil test index to residual or nutrient carryover, the latter being in units equivalent to those of applied fertilizer nutrient. The soil test index of the preceding season is given by $\mathrm{b}_{\mathrm{it}-1}^{\star}+\lambda_{i s}^{-1} \mathrm{X}_{\text {it-1 }}$ where $\mathrm{X}_{i t-1}$ is fertilizer applied at the beginning of the season. A proportion of that index is made available to the crop. This season's soil test should therefore be expected to be related to that of the previous season in a manner dictated by site and weather variables unique to that location as follows:

$$
\begin{equation*}
b_{i t}^{\star}=g\left(b_{i t-1}^{*}+\lambda_{i s}^{-1} x_{i t-1}\right) \tag{3.8}
\end{equation*}
$$

where $\lambda_{\text {is }}^{-1} x_{i t-1}$ is in $\mu g / g$, all the symbols having been explained in the text. The inverse of (3.8) with variables measured in kg of applied fertilizer is

$$
\begin{equation*}
b_{i t}=h\left(b_{i t-1}+x_{i t-1}\right) . \tag{3.9}
\end{equation*}
$$

It should be noted that equation (3.9) does not involve soil tests. The relationship $b_{i}=\lambda_{i} b_{i}^{*}$ in (3.6) implies that $b_{i}^{*}=\lambda_{i}^{-1} b_{j}$ and since $b_{i}$ is measured in units equivalent to those of currently applied fertilizer $X_{i t}$, the latter can similarly be converted to soil test units by multiplying by $\lambda_{i}^{-1}$ thus explaining the reciprocal relationship between equations (3.8) and (3.9).

Tracing the destination of nutrients, particularly the relatively mobile nutrients such as nitrogen by means of chemical analysis of soils can be a difficult task as noted by Stauber and Burt [1973]. The authors suggested an implicit way of measuring nutrient carryover. An extra variable was introduced in $h(\cdot)$ and the function specified as

$$
\begin{aligned}
& b_{N t}=\alpha_{0}\left(b_{N t-1}+x_{N t-1}\right)^{\alpha_{1}} W_{t-1}^{\alpha_{2}} \\
& \text { where } W_{t}=\text { precipitation during growing season } \\
& b_{N t}=\text { nitrogen carryover } \\
& \text { and } \quad \alpha_{i}=\text { parameters; } i=0,1,2 .
\end{aligned}
$$

The authors were interested in explaining the response of grass to total nitrogen by employing the function

$$
\begin{aligned}
& Y_{t}=f\left(X_{N t}^{\top}, W_{t}, \beta\right) \\
& \text { where } X_{N t}^{\top}=\underset{\text { total supply of nitrogen: applied }\left(X_{N t}^{a}\right) \text { plus }}{\text { carryover }\left(b_{N t}\right)} \text { ) } \\
& \text { and } \beta=\text { parameter vector. }
\end{aligned}
$$

The parameter vectors, $\alpha$ of the carryover function and $\beta$ of the yield function, are estimated simultaneously after substituting for ${ }^{b}{ }_{N t}\left(X_{N 1}^{a} \ldots X_{N t}^{a}\right)$ in (3.11). Details of the procedures and its pitfalls can be found in Stauber and Burt [1973] and Stauber et al., [1975].

The major attributes of the author's approach are that no soil tests are required, a weather variable is introduced to explain variations in nutrient carryover and finally, fertilizer recommendations on the basis of their approach requires knowledge of past fertilizer applications. The latter attribute appears restrictive but as Matocha et al., [1970] have noted, past fertilizer applications may be a necessary requisite in recommendations of fertilizer use whenever it appears that the soil
tests are not correlated well with actual nutrient release. This may very well be the case for nitrogen, a nutrient which is known to occur in highly transient forms in many soil-type and weather combinations. ${ }^{1 /}$

Other authors have noted the dependence of nutrient carryover on weather variables and crop species. Different crops have different capabilities to extract native nutrients as noted for example by Fixen and Carson [1978]. Weather variables such as temperature and precipitation influence the mobility and hence availability of the elements in the soil. However, only a few authors have actually formulated nutrient carryover explicitly as a function of these variables. Fuller [1965, p. 111] attempted to explain the complexity of the relationship between past yields, weather variables, and nitrogen carryover. The author employed a model in which total nitrogen $N_{t}^{\top}$ is given by $N_{t}^{a}+C_{t}+U_{t} . C_{t}$ is the nitrogen carryover and $U_{t}$ is a stochastic term which is a function of rainfall and past yields. This formulation makes $N_{t}^{\top}$ stochastic and hence the crop response too is stochastic.

The foregoing discussion points to the fact that past yields should be included in the carryover function in (3.9). In many circumstances the harvested portion of the crop represents a permanent loss of nutrients from the soil. Past yields should therefore be considered as a proxy to the actual elemental quantities extracted from the soil by way of harvesting. If the harvested portion of the crop does not represent a significant nutrient loss, the inclusion of past yields in the carryover function cannot be defended.

[^4]Weather and site characteristics such as soil pH, temperature, rainfall, and past management and their interactions may also explain some of the variability in nutrient carryover [Fixen and Carson, 1978, Fuller, 1965].- The management aspect is possibly a major factor determining loss of soil nutrient through soil erosion and run-off. But, the picture is further complicated by the fact that, depending on the nutrient mobility, which in turn is a function of the chemical form in which the nutrient exists in a particular soil, significant quantities can be lost by leaching into aquifers, rivers, and lakes. The use of data from experiments in which standard field management practices are adhered to may rationalize the exclusion of some of the site variables. In some cases the variability in soil pH is restricted within prescribed bounds by addition of specified amounts of lime to the soil. This would make pH an unimportant explanatory variable in the carryover function in (3.9). It is, however, not feasible to control all the variables even under experimental conditions. The relationship between $b_{i t}$ and the included variables is, therefore, not deterministic. To account for this, a stochastic error term with specified properties should be included in the relationship.

Where the soil tests are reflective of the exchangeable quantities of soil nutrients, equation (3.8) may be preferrable to (3.9) on grounds that the variable $b_{i t}$ in (3.9) is not observable. As already noted, there are cases when (3.9) may be taken in favor of (3.8), the attendant complications involved in implicit estimation of $b_{i t}$ notwithstanding. In this particular study, however, the carryover formulation in terms of soiltest units will be adopted. In light of the foregoing discussion on variables affecting nutrient carryover, equation (3.8) is now written
in the form

$$
\begin{align*}
b_{i t}^{*}=g\left(b_{i t-1}^{*}+\lambda_{i s}^{-1} x_{i t-1}^{a}, y_{t-1}\right)+v_{i t}  \tag{3.12}\\
\text { where } y_{t-1}=\text { crop yield prior to soil sampling at period } t \\
v_{i t}=\text { stochastic error term }
\end{align*}
$$ and other symbols are as previously defined.

Equation (3.12) relates the current index for the $i^{\text {th }}$ nutrient availability, $b_{i t}^{*}$, with the index of total availability at the beginning of the preceding period and crop yield prior to soil testing. Where nutrient losses through leaching are not significant, or alternatively, for relatively immobile nutrients such as phosphorous and potassium, crop harvests may be a major source of nutrient loss. Viewing the variable $y_{t-1}$ as a nutrient extractor in conjunction with the fact that plant roots have to forage for the relatively immobile nutrients [Bray, 1954] leads to the conclusion that the direction of causality is from $y_{t-1}$ to $b_{i t}^{*}$. This in turn implies that a bumper crop should be followed by a low soil test and vice versa. But this may be true only if the harvested portion of the crop has a high nutrient content compared to the portion which is recycled back into the field. The relationship between $y_{t-1}$ and $b_{i t}^{*}$ may also be confounded by the influences of other unspecified variables such as temperature and moisture content of the soil. It is possible,' therefore, that the preceding harvest may be correlated with the current soil test for reasons other than that the former is an avenue of nutrient loss. A different perspective to the same problem is by viewing the carryover $b_{i t}^{*}$ as a function of all past events which also condition the yield $y_{t-1}$. This reinforces the notion just mentioned above, namely that $\mathrm{b}_{\mathrm{it}}^{\star}$ and $y_{\mathrm{t}-1}$ may be correlated even when the nutrient loss embodied in $y_{t-1}$ is insignificant.

In the special case of a rotation in which soil sampling is conducted only once during a rotation, as will be encountered in Section 3.3, y $\mathrm{t}_{\mathrm{t}-1}$ in (3.12) would be a vector of all crop yields in the preceding rotation. This would be a rational and unquestionable procedure if the direction of causality was known to be from previous crop yield to current soil test. However, in light of the uncertainties and the possible relationships mentioned in the paragraphs above, the question of which particular crop(s) to be obtained in (3.12) on grounds of variation in $b_{i t}^{\star}$ explained can only be settled inempirical terms.

In Section 3.1, equation (3.5) was simply expressed in terms of total available nutrients, $x_{i}^{\top}$. By definition, $x_{j}^{\top}=b_{i}+x_{i}$ where $b_{i}$ is nutrient carryover expressed in kg of applied fertilizer nutrients. As already noted, $b_{i}$ is not observable. It is, however, estimated by $\lambda_{i s}{ }^{b_{i}^{*}}$ where $b_{i}^{*}$ is in soil-test units and $\lambda_{i s}$ is estimated from an equation such as (3.12). By substituting for $b_{i}$ in equation (3.5) of Section 3.1, the crop response equation becomes

$$
\begin{equation*}
y_{t}=A_{s w_{t}} \operatorname{Min}\left\{f\left(\lambda_{i s} b_{i t}^{\star}+x_{i t}\right)\right\}+\varepsilon_{t} \tag{3.13}
\end{equation*}
$$

where all the symbols have the same interpretation as in the preceding equations.

Equation (3.13) and the carryover function in (3.12) constitute a dynamic non-substitution crop response model. The term non-substitution has already been explained in Section 3.1. The dynamic aspect of the model derives from the fact that current economic optimization over $X_{i}$ is conditioned by nutrient carryover. The latter is a function of past actions and events.

The model has two equations and two endogenous variables $y$ and $b_{i}^{*}$. The appropriate technique for estimating the two equations is therefore
that which is applicable to a simultaneous-equations system. The endogenous variable y does not appear in the carryover equation. In addition, if it is assumed that the contemporaneous covariances are zero implying that $\operatorname{Cov}\left(\varepsilon, V_{i}\right)=\Sigma$ is diagonal, the model would qualify as a recursive system [Intrilligator, 1978, p. 359]. Equations of a recursive system can be estimated individually without risking biased estimates of the parameters.

A description of the stochastic properties of the error term, $\mathrm{v}_{\mathrm{it}}$, will not be attempted at this juncture. That task is taken up in Section 3.3 which deals with the interpretation and estimation techniques for the carryover function in (3.12). Suffice it here to mention that the error term $\varepsilon_{t}$ in (3.13) embodies omitted variables and measurement errors relevant to the growing season. Soil sampling is conducted at the beginning of the season, thus, it may not be unrealistic to assume that the random events implicit in $v_{i t}$ are independent of those in $\varepsilon_{t}$ for all time periods. The estimation of equations (3.12) and (3.13) will therefore proceed under the premise that the model is recursive. Since the $\lambda$-values are needed in equation (3.13), the first priority will be given to the carryover equation. Once the estimate of $\lambda_{\text {is }}$ is obtained, a discussion of the estimation techniques for the crop response function can then proceed in terms of $\hat{X}_{i}^{\top}$ which is given by $\hat{\lambda}_{i s} b_{i t}^{\star}+x_{i t}$. The first part of the next section will, however, be devoted to a description of the data to be used for estimation.

### 3.3 Estimation of the carryover functions

### 3.3.1 The data ${ }^{1 /}$

The experiment was started in 1952 on the Agronomy Farm at Purdue University on a Raub silt loam, an imperfectly drained prairie soil. The crops initially used were corn (Zea mays L.), soybean (Glycine max L.), wheat (Triticum vulgare L.) and hay (a mixture of alfalfa (Medicago sativa L.), red clover (Trifolium pratense L.), and brome grass (Bromus sp.)). The applied nutrients were phosphorous ( $P$ ) and potassium (K) in the form of superphosphate and potassium chloride respectively.

There were 22 treatments of $P$ and $K$ randomized within each of the 8 blocks representing two replicates. Crops were randomized within each replicate. The rotation sequence was corn-1-soybean-wheat-hay. In 1963 the hay crop was replaced by a second crop of corn designated here as corn-2.

All the straw was ploughed back into the field and soil samples were taken only from the hay plots prior to broadcasting $P$ and $K$ in fall. Corn-1 was planted on these same plots in May of the following year. This means that for a given block, both soil sampling and broadcasting of $P$ and $K$ took place only once in four years. Soybean, wheat, and hay did not receive direct broadcast applications of $P$ and K. Only corn-1 and wheat received row applications of fertilizer. Nitrogen was plowed under for corn-1 plots and top-dressed to the wheat plots at non-limiting levels.

[^5]The soil pH was adjusted to 6.5 by applying suitable amounts of lime.
Experimentation continued through 1980. The soil sampling and fertilization procedures are presented schematically on Table 3.1. Application rates for $P$ and $K$ are presented in Appendix Table A.2.

Table 3.1 is only a stylization of the operations involved in the experiment. Some essential details such as how and specifically when fertilizer was broadcast have been omitted. The important fact to remember is the sequence in the rotation as it relates to soil sampling and the fertilization routines. This fact is crucial to the development and interpretation of subsequent mathematical formulations.

The soil sampling procedure involved taking 15 cores in the plow layer, $0-15 \mathrm{~cm}$, from the central portion of each of the hay plots. The plots measured 4.3 m wide and 19.8 m long.

The available phosphorous and potassium was extracted at the Purdue Soil Testing Laboratory by shaking 5 g of soil with 15 ml of 0.7 HCl in a shaker for two minutes. In 1968, the phosphorous extraction procedure was changed to Bray Pl. At a soil pH of about 5.8, the average initial soil tests, before $P$ and $K$ applications in 1952, were 18 and $45 \mu \mathrm{~g} / \mathrm{g}$ respectively.

### 3.3.2 A distributed lag model interpretation of the carryover functions

On the basis of the foregoing account fo the experimental design and field operations summarized in Table 3.1, the following carryover is proposed

$$
\begin{align*}
& b_{i j k t}^{*}=g_{i}\left[b_{i j-1 k t}^{*}+\lambda_{i}^{-1}\left(x_{i j k-1 t}^{R}+x_{i j k-3 t}^{R}+x_{i j-1 k t}^{B}\right), y_{j k t}^{H},\right. \\
&\left.y_{j k-1 t}^{W}, y_{j k-2 t}^{S}, y_{j k-3 t}^{C-1}\right]+V_{i j k t} \tag{3.14}
\end{align*}
$$

Table 3.1 Summary of Field Operations at the Purdue Experiment

| ROTATION (J) | 1 | 2 | 3 | 4 | 5 | . . . |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| CALENDAR YEAR | 5253545556 | 57585960 | 61626364 | 65666768 | 69707172 | . . |
| PERIOD IN ROTATION (K) | $0 \quad 1 \quad 2 \quad 3$ | 1234 | $1 \begin{array}{llll}1 & 2 & 4\end{array}$ | 1234 | $1 \begin{array}{llll}1 & 2 & 3 & 4\end{array}$ | . . |
| BLOCK I | $\begin{array}{rrrrr} B & R & & R & B \\ C 1 & S & W & H \\ b_{0}^{*} & & & & b_{14}^{\star} \end{array}$ | $\begin{array}{cccc} R & & R & B \\ C 1 & S & W & H \\ & & & b_{24}^{\star} 4 \end{array}$ | $\begin{array}{rrrr} R & & R & B \\ C 1 & S & W & C 2 \\ & & & b_{34}^{\star} \end{array}$ | $\begin{array}{rlll} \mathrm{R} & & \mathrm{R} & \mathrm{~B} \\ \mathrm{C} 1 & \mathrm{~S} & \mathrm{~W} & \mathrm{C} 2 \\ & & & \mathrm{~b}_{44}^{*} \end{array}$ | $\begin{array}{rlll} \mathrm{R} & & \mathrm{R} & \mathrm{~B} \\ \mathrm{C} 1 & \mathrm{~S} & \mathrm{~W} & \mathrm{C2} \\ & & & \mathrm{~b}_{54}^{\star} \end{array}$ | - • |
| II | $b_{0}^{\star} b_{11}^{\star}$ | $\begin{array}{ccc} \mathrm{HCl} & \mathrm{~S} & \mathrm{~W} \\ \mathrm{~b}_{21}^{*} & & \end{array}$ | $\begin{array}{lll} \mathrm{HCl} & \mathrm{~S} & \mathrm{~W} \\ \mathrm{~b}_{31}^{\star} & & \end{array}$ | $\begin{aligned} & \mathrm{C} 2 \mathrm{Cl} \quad \mathrm{~S} \\ & \mathrm{~b}_{41}^{\star} \end{aligned}$ | $\begin{array}{llll} \mathrm{C} 2 \mathrm{C} 1 & \mathrm{~S} & \mathrm{~W} \\ \mathrm{~b}_{51}^{*} & & \end{array}$ | -•• |
| II I | $\begin{array}{lll}  & \text { W H Cl } & \text { S } \\ \mathrm{b}_{0}^{\star} & \mathrm{b}_{12}^{\star} \end{array}$ | $\begin{gathered} \text { W } \underset{\substack{\text { H Cl } \\ b_{22}^{\star}}}{\text { S }}=2 \end{gathered}$ | $\begin{gathered} \text { W H C1 S } \\ \mathrm{b}_{32}^{\star} \end{gathered}$ | $\begin{gathered} \text { WC2 C1 } \mathrm{S} \\ \mathrm{~b}_{42}^{\star} \end{gathered}$ | $\begin{gathered} \text { WC2 C1 } \mathrm{S} \\ \mathrm{~b}_{5}^{\star} 2 \end{gathered}$ | - . . |
| IV | $\begin{array}{ccc}  & \text { S W H C1 } \\ \mathrm{b}_{0}^{\star} & \mathrm{b}_{13}^{\star} \end{array}$ | $\begin{gathered} S \quad W \mathrm{HC1} \\ b_{23}^{\star} \end{gathered}$ | $\begin{gathered} S \quad W \quad H C 1 \\ b_{33}^{*} \end{gathered}$ | $\begin{gathered} S \quad \text { W C2 C1 } \\ b_{43}^{\star} \end{gathered}$ | $\begin{gathered} \mathrm{S} \quad \mathrm{~W} \mathrm{C} 2 \mathrm{Cl} \\ \mathrm{~b}_{53}^{\star} \end{gathered}$ | -•• |

$R=$ Row application of $P$ and $K$ shown here only for Block $I$ in order to avoid clutter.
$B=$ Broadcast application of $P$ and $K$ done after soil sampling from the hay plots, shown here only for Block I.
$b_{j k}^{*}=$ Soil sampling, after harvesting hay/corn-2; $b_{o}^{*}$ represents initial soil sampling.
Other letters represent the four crops: corn-1, soy, wheat, hay and corn-2.

$$
\begin{aligned}
& \text { where } b_{i j k t}^{*}=\text { soil test unit of the } j^{\text {th }} \text { nutrient at the } t^{\text {th }} \\
& \text { treatment in the } k^{\text {th }} \text { period of the } j \text { th rotation } \\
& x_{i j k-1 t}^{R}+X_{i j k-3 t}^{R}=\text { total row application of the } i^{\text {th }} \text { nutrient at } \\
& \text { the } t^{\text {th }} \text { treatment during the } j^{\text {th }} \text { rotation; } 1 / 4 \\
& \text { of the fertilizer is applied to wheat one period } \\
& \text { prior to soil sampling and } 3 / 4 \text { to corn-1 three } \\
& \text { periods back } \\
& \begin{aligned}
x_{i j-1 k t}^{B}= & \text { broadcast application of the } i^{\text {th }} \text { nutrient at } \\
& \text { the th treatment in period } k \text { of the previous } \\
& \text { rotation }
\end{aligned} \\
& \lambda_{i}=\begin{array}{l}
\text { an unknown constant of proportionality for the } \\
\\
i \text { th nutrient; only one soil type is involved }
\end{array} \\
& \text { hence the subscript s introduced in Section } 3.2 \\
& \text { is dropped } \\
& y^{\mathrm{H}}, y^{\mathrm{W}}, y^{\mathrm{S}}, y^{\mathrm{Cl}}=\text { yield of hay, wheat, soybean, and corn-1 in that } \\
& \text { order. From } 1963 y^{7} \text { is substituted by } y C 2 \text {, yield } \\
& \text { of corn-2 }
\end{aligned}
$$

and $\quad V_{i j k t}=\begin{aligned} & \text { an error term whose stochastic properties will } \\ & \\ & \text { be defined later. }\end{aligned}$
By letting $r$ represent replication, a more compact form of equation
(3.14) for the $i^{\text {th }}$ nutrient and $j^{\text {th }}$ rotation is given by

$$
\begin{aligned}
& B_{i j}^{*}=G_{i}\left[B_{i j-1}^{*}+\lambda_{i}^{-1} X_{i j-1}^{A}, Y_{j}\right]+V_{i j} \\
& \text { where } B_{i j}^{*}=\left[\begin{array}{c}
B_{i j, 1}^{*} \\
\hdashline B_{i j, 2}^{*}
\end{array}\right] \\
& \text { and } B_{i j, r}^{*}=\left[\begin{array}{c}
b_{i j 101, r}^{*} \\
\vdots \\
b_{i j}^{*} \\
\vdots \\
b_{i j}^{*} 122, r \\
\hline
\end{array}\right]
\end{aligned}
$$

The other variables in (3.15) are constructed in a manner similar to the construction of $B_{i j}^{*}$ shown above where $i=1,2 ; j=1,7 ; k=1,4$ and $t=1,22 . x_{i j-1}^{A}$ is the total nutrient application during the outgoing
rotation. It should be noted from equation (3.14) that strictly, the subscript j-1 applies only to the broadcast component of $x^{A}$, the other being the row application to corn-1 and wheat of the current rotation. The construction of the column vectors in (3.15) must therefore be done with close reference to equation (3.14).

The lag structure in (3.15) is with respect to both the rotations and the periods within a rotation. The vector of crop yields $Y_{j}$ for instance represents events taking place before the soil testing is done. Remember that soil sampling is done after harvesting hay/corn-2 but prior to broadcasting $P$ and $K$ for corn-1. The variables under $G_{i}[\cdot]$ in equation (3.15) are indicators of the history of the plot from which the current soil sample is taken. In words, therefore, equation (3.15) says that the current rotation soil test for the $i^{\text {th }}$ nutrient at any given plot depends on the total supply of that nutrient (given by the index $B_{i j-1}^{*}+\lambda_{i} X_{i j-1}^{A}$ ) and the amount of nutrient loss (given by $Y_{j}$ ) during the outgoing rotation. The latter symbol is an $N \times 4$ matrix whose $t^{\text {th }}$ row, for given $j$ and $k$, is $\left(y_{t}^{H}, y_{t}^{W}, y_{t}^{S}, y_{t}^{C 1}\right) . N$ is the total number of observations equal to $J \times K \times T$ where the latter symbols represent totals of rotations, periods, and treatments respectively.

The sequence of rotation was C1-S-W-H(C2) soil sampling was done after harvesting hay. As already stated, only the hay plots were sampled. It is therefore not unreasonable to expect yields far removed from the period of soil sampling to have a less significant role in equation (3.15). Since crop harvests represent nutrient loss, all the coefficients of the components of $Y_{j}$ should be expected to be negative. Because of complications arising from stochastic relationships, this need not always be the case. All the four crops will therefore be retained in (3.15) and
exclusion of variables done only on the basis of preliminary empirical results.

The influence of $X_{i j-1}^{A}$ and $B_{i j-1}^{*}$ on $B_{i j}^{*}$ is rather straightforward and will therefore not be discussed here.

The use of a common constant of proportionality $\left(\lambda_{i}\right)$ in (3.15) to relate soil tests to fertilizer applied at different periods during a rotation is questionable. Referring back to equation (3.14), let $X_{i j-1 k t}^{C}=x_{i j k-3 t}^{R}+X_{i j-1 k t}^{B}$ be the total fertilizer applied to corn-1. Notice the simplification in notation; $X_{i j-1 k t}^{B}$ is fertilizer broadcast immediately after soil sampling, four periods back, while $X_{i, j k-3 t}^{R}$ is row application almost half a period later. $X_{i j k-3 t}^{R}$ as a proportion of $X_{i j-1 k t}^{C}$ is insignificant, hence the use of the reconstructed variable ( $X_{i j-1 k t}^{C}$ ) should not lead to serious biases in estimation. $x_{i j k-1 t}^{R}$, on the other hand, is the row application of the $i^{\text {th }}$ fertilizer nutrient to wheat, an event which in relation to soil sampling is two periods back (see Table 3.1).

Let fertilizer application to wheat be denoted by $X_{i j k-1 t}^{W}$. An alternative carryover function, written in compact notation similar to that in (3.15), is given as

$$
\begin{equation*}
B_{i j}^{*}=G_{i}\left[B_{i j-1}^{*}+\lambda_{i 4}^{-1} X_{i j-1}^{C}+\lambda_{i 2}^{-1} X_{i j}^{W}, Y_{j}\right]+V_{i j} \tag{3.16}
\end{equation*}
$$

in which current soil test is related to soil test, fertilizer nutrient applied to corn-1 and wheat and yields of the four crops in the outgoing rotation. As before, $V_{i j}$ is an error term.

Equation (3.15) differs from (3.16) in only one respect, namely the use of two different constants of proportionality, $\lambda_{i 4}$ and $\lambda_{i 2}$, in the latter. The first constant refers to fertilizer applied (approximately) four periods back while the second refers to fertilizer applied only
two periods back. For a given soil-type $\lambda_{i}$ has been defined in Section 3.2 as the proportion of the stock of the $i^{\text {th }}$ nutrient in the soil released during the growing season. It follows that these proportions depend on length of the growing season and that in equation (3.16), $\lambda_{i 2}^{-1} \geqq \lambda_{i 4}^{-1}$. In fact, since $\lambda_{i 4}$ refers to fertilizer applied four periods back and $\lambda_{i 2}$ to fertilizer applied two periods before soil sampling, the value of $\lambda_{i 4}^{-1}$ may be approximately half that of $\lambda_{i 2}^{-1}$, assuming a proportional relationship. Using a similar reasoning, one may obtain a seasonal (one period) $\lambda$-value for the $i^{\text {th }}$ nutrient by extrapolation. It is these seasonal $\lambda$-values which will be used for constructing the total seasonal nutrient supply given by $\hat{X}_{i t}^{\top}=\hat{\lambda}_{i} b_{i t}^{\star}+x_{i t}$ as explained in Section 3.2.

Equation (3.16) will be interpreted as a reduced form of a distributed lag model [see e.g., Lanzer and Paris, 1981]. A rotation is the basic time framework but the cross-sectional aspect of the equation should be kept in mind. Estimation of the equation requires a specification of the stochastic properties of the error term $V_{i j}$. First, the equation is written explicitly as:

$$
\begin{aligned}
& B_{i j}^{\star}=\gamma_{i}\left[B_{i j-1}^{\star}+\lambda_{i 4}^{-1} X_{i j-1}^{C}+\lambda_{i 2}^{-1} X_{i j}^{W}\right]+Y_{j} \alpha_{i}+V_{i j} \\
& \text { where } \quad \gamma_{i}= \text { geometric decline coefficient whose absolute value } \\
& \text { is assumed to be less than one }
\end{aligned} \quad \begin{aligned}
\alpha_{i}= & \begin{array}{l}
\text { coefficient vector conformable to } Y_{j} \text { which is the } \\
\\
\text { vector of crop yields in the } j \text { th rotation. }
\end{array}
\end{aligned}
$$

and other symbols retain the interpretation already given.
A lag operator (L) will be introduced so as to simplify the notation in the equations that follow. For any given variable, $X, L X_{j}=X_{j-1}$, $L^{2} X_{j}=X_{j-2}$ and in general $L S_{j}=X_{j-S}$. Also, $\sum_{S=0}^{\infty} \gamma_{L} S_{X_{j}}=1 /(1-\gamma L)$ given that $|\gamma|<1$.

Equation (3.17) is the reduced form of the following distributed lag model:

$$
\begin{align*}
B_{i j}^{\star}= & \frac{1}{1-\gamma_{i} L}\left[\gamma_{i} \lambda_{i 4}^{-1} x_{i j-1}^{C}+\gamma_{i} \lambda_{i 2}^{-1} x_{i j}^{W}+Y_{j} \alpha_{i}\right]+U_{i j} \\
= & \sum_{S=0}^{J-1} \gamma_{i} L^{S}\left[\gamma_{i} \lambda_{i 4}^{-1} x_{i j-1}^{C}+\gamma_{i} \lambda_{i 2}^{-1} x_{i j}^{W}+Y_{j} \alpha_{i}\right] \\
& +\sum_{S=j}^{\infty} \gamma_{i}^{S} S^{S}\left[\gamma_{i} \lambda_{i 4}^{-1} x_{i j-1}^{C}+\gamma_{i} \lambda_{i 2}^{-1} x^{W}+Y_{j} \alpha_{i}\right] \\
& +U_{i j} ; j=1, J . \tag{3.18}
\end{align*}
$$

The term summing to infinity ( $\infty$ ) on the right hand side of (3.18) is often referred to as the truncation remainder. It can be summarized by $\eta_{j} \gamma_{j}^{j}$ where $\eta_{i}$ is a function of the missing information, $S=(J, \infty)$. Equation (3.18) in the form

$$
\begin{equation*}
B_{i j}^{\star}=\sum_{S=J}^{J-1} \gamma_{i} L^{S}[\cdots]+n_{i} \gamma_{i}^{j}+U_{i j} \tag{3.19}
\end{equation*}
$$

is a result of the estimation procedure proposed by Klein [1958]. Maddala [1977, pp. 360-362] has a brief discussion of an iterative procedure for estimating it directly.

As it turns out, however, the truncation remainder (or the missing information) problem is a small sample size problem. As the sample size becomes infinite, the contribution of the truncation remainder diminishes, given that $\left|\gamma_{i}\right|<1$. Under such circumstances, the values chosen for $\eta_{i}$ are inconsequential for the asymptotic properties of the other parameters in (3.19) as discussed by Dhrymes et al. [1970] and Dhrymes [1971]. For large samples, the second term on the right hand side of (3.19) is dropped giving a carryover function of the form

$$
\begin{equation*}
B_{i j}^{*}=\frac{1}{1-\gamma_{j} L}[\cdots]+U_{i j} \tag{3.20}
\end{equation*}
$$

where the terms enclosed in the square brackets are those given in equation (3.18). It must be emphasized that the step leading from (3.19) to (3.20) is justified only if the sample size is considered as large. The nature of the moving rotation and cross-sectional structure of the Purdue long term experiment is such that it is not unreasonable to operate under the premise of a large sample size. Subsequently, a large sample size will be presumed.

The reduced form equation in (3.17) is obtained from (3.20) by multiplying the latter through by $1-\gamma_{j} L$. This implies that the error term in (3.17) is related to that in (3.20) as follows:

$$
\begin{equation*}
v_{i j}=U_{i j}-\gamma_{i} U_{i j-1} \tag{3.21}
\end{equation*}
$$

The basic differences among the techniques used for estimating distributed lag models derive from the assumptions made regarding the stochastic properties of the disturbance term such as $U_{i j}$ in (3.20). Equation (3.20) can be estimated directly by use of ordinary least squares (OLS) as originally proposed by Klein [1958] and further discussed by Pesaran [1973] and Maddala [1977]. The common practice of estimating the reduced form given in (3.17) will be adopted here. Details concerning the various estimation techniques can be found for example in Dhrymes [1971].

First, consider the case when the stochastic components of $U_{i j}$ are assumed to be independently and identically distributed (iid). In keeping with the notation adopted earlier, the components of $U_{i j}$ are $U_{i j k t, r}$ referring to the disturbance term at the $t^{\text {th }}$ treatment in the $k^{\text {th }}$ period of the $j^{\text {th }}$ rotation. The other subscripts $i$ and $r$ refer to nutrient and replication respectively. It is also common practice to assume that $E\left(U_{i j}\right)=0$ and that for a given replication, the terms $U_{i j k t}$ have a
common variance $\sigma_{r}^{2}$.
The assumptions given above imply that $\operatorname{cov}\left(U_{i j}\right)=E\left(U_{i j} U_{i j}^{\prime}\right)=\Sigma$ where the latter is a block diagonal matrix with blocks given by $\sigma_{r}^{2} I$; $r=1,2$. The Purdue experiment was conducted at the same geographic location under similar field management practices in the two replications. Hence, there seems to be no reason to assume that $\sigma_{1}^{2} \neq \sigma_{2}^{2} . \operatorname{Cov}\left(U_{i j}\right)$ will therefore be taken as $\sigma^{2} I$ where $\sigma^{2}=\sigma_{1}^{2}=\sigma_{2}^{2}$ and $U_{i j}=\left(U_{i j, 1}, U_{i j, 2}\right)^{\prime}$.

Given that the elements of $U_{i j}$ are iid as assumed above, the corresponding elements of $\mathrm{V}_{1 j}$ in the reduced form equation, (3.17), will be autoregressive as can be deduced from (3.21). Under such circumstances, efficient estimation of the parameters of the reduced form equation requires prior knowledge of $\Omega^{-1}$, the inverse of $\operatorname{cov}\left(V_{i j}\right)$. Applying OLS to appropriately transformed data (using $\Omega^{-1}$ ) is equivalent to maximum likelihood estimation if the assumptions are true.

If $\Omega^{-1}$ is unknown, consistent estimates of the parameters in (3.17) can be obtained by applying two-step procedures in conjunction with instrumental variables estimation to provide starting values. Alternatively, one can employ various modifications of generalized least squares as discussed by Dhrymes [1971].

Estimation of (3.17) under the premise that $\operatorname{cov}\left(U_{i j}\right)=\sigma^{2} I$ may not be realistic in the case of the long term rotation experiment being considered here. At this point, it may be helpful to refer back to Table 3.1. The error term $U_{i j k}$ for a given rotation ( $j$ ) and nutrient ( $i$ ) is a vector of residual terms from different blocks since soil samples were taken only from hay blocks. For a given rotation, therefore, $U_{i j k}$ for $k=1,4$ are independent vectors. Within period $k$, the error terms $U_{i j k t}$ which constitute the vector $U_{i j k}$ can be considered as being generated
randomly and independently; t refers to treatment levels. Such errors may be due to measurement during soil sampling and testing.

When one moves from one rotation to the next, the independence assumption made above for within rotation residuals may not be valid. As can be seen from Table 3.1 , the soil test symbol $b_{j k}^{*}$ falls on the same block every four years. This is also true for fertilizer applications and the entire cropping pattern. Hence, it is only the events between rotations which refer to the same blocks and treatment levels. Any stochastic relationship between residuals, if it exists, must therefore be sought only between residuals of different rotations. Thus, within a rotation, the residuals are cross-sectional in nature and are assumed to be independent. From rotation to rotation, the residuals are given a time series interpretation. It is being assumed here that the relationship between the current residual and those in past rotations follows a simple Markov process which will be written compactly as:

$$
\begin{align*}
& U_{i j}=\frac{1}{1-\rho_{i} L} W_{i j}  \tag{3.22}\\
& \text { where } \quad \rho_{i}=\text { autoregressive parameter, }\left|\rho_{i}\right|<1 \\
& L=\text { lag operator }
\end{align*}
$$

$$
\text { and } W_{i j} n, N\left(0, \sigma^{2} I\right)
$$

The relationship in (3.20) now becomes

$$
\begin{aligned}
B_{i j j}^{\star}= & \frac{1}{1-\gamma_{i} L}[\cdots]+\frac{1}{1-\rho_{i} L} W_{i j} \\
\text { where } B_{i, j}^{\star} & =\text { vector of soil tests in } \mu g / g \\
\gamma_{i} & =\text { geometric decline parameter, }\left|\gamma_{i}\right|<1 \\
W_{i j} & =\text { white noise with the distribution given above } \\
\text { and }[\cdots] & =\text { terms in square brackets in (3.18). }
\end{aligned}
$$

The reduced form given in (3.17) is obtained from (3.23) by first substituting $U_{i j}$ for the last term in the latter and multiplying through by
$1-\gamma_{j} \mathrm{~L}$.
In the special (rather trivial) case that $\gamma_{i}=\rho_{i}$, applying OLS directly to (3.23) will yield consistent estimates of the relevant parameters; since then, $V_{i j}$ is equivalent to $W_{i j}$. If $\gamma_{i}$ is not equal to $\rho_{i}$, the estimation of (3.23) by OLS does not generally give consistent estimates of the parameters. However, the use of OLS in conjunction with numerical techniques is equivalent to MLE if applied to the following model derived from (3.23) by multiplying through, first by $1-\gamma_{j} L$ and then by $1-\rho_{i} L$ :

$$
\begin{align*}
B_{i j}^{\star}= & \rho_{i} B^{*}{ }_{i j-1}+\gamma_{i}\left(B_{i j-1}-\rho_{i} B_{i j-2}\right)+\gamma_{i} \lambda_{i 4}^{-1}\left(X_{i j-1}^{C}-\rho_{i} X_{i j-2}^{C}\right) \\
& +\gamma_{i} \lambda_{i 2}^{-1}\left(X_{i j}^{W}-\rho_{i} X_{i j-1}^{W}\right)+\alpha_{i 1}\left(Y_{j}^{H}-\rho_{i} Y_{j-1}^{H}\right) \\
& +\alpha_{i 2}\left(Y_{j}^{W}-\rho_{i} Y_{j-1}^{W}\right)+\alpha_{i 3}\left(Y_{j}^{S}-\rho_{i} Y_{j-1}^{S}\right) \\
& +\alpha_{i 4}\left(Y_{j}^{C 1}-\rho_{i} Y_{j-1}^{C 1}\right)+W_{i j} \tag{3.24}
\end{align*}
$$

where all the symbols have been explained in the preceding equations and, as before, $W_{i j} \sim N\left(0, \sigma^{2} I\right)$.

Equation (3.24) can be estimated iteratively by use of OLS as explained by Theil [1971, pp. 422-425]. Because $W_{i j}$ is assumed to be white noise entering (3.24) additively, such a procedure would lead to estimators with desirable asymptotic properties. In the present case, however, equation (3.24) is estimated directly by a nonlinear least squares technique referred to as SHAZAM and developed by White [1978].

### 3.4 Estimation of the nutrient non-substitution model

### 3.4.1 Alternative approaches

In Section 3.2 an assumption was made to the effect that the error terms $\varepsilon_{t}$ and $V_{t}$ are uncorrelated. This allowed the carryover functions to be estimated individually. As already noted, the estimate of $\lambda_{i}$ is used in the nonsubstitution model given in equation 3.5 (Section 3.1). The total supply of the $i^{\text {th }}$ nutrient, $X_{j}^{\top}$, is then estimated by $\hat{X}_{i}^{\top}=\hat{\lambda}_{i s} b_{i}^{\star}+X_{i}$ where $\lambda_{i s}$ is an unknown constant of proportionality for a given soil type, $b^{*}$ is the soil test in $\mu \mathrm{g} / \mathrm{g}$ and X is the applied fertilizer nutrient.

The set of nutrients will be taken as two. For a given time framework ( $t=1, N$ ), equation (3.5) is written as

$$
\begin{equation*}
y_{t}=A \operatorname{Min}\left\{f_{1 t}, f_{2 t}\right\}+\varepsilon_{t} \tag{3.25}
\end{equation*}
$$

For simplicity, the $S$ and $W$ subscripts on the maximum yield ( $A$ ) have been dropped and $f_{i t}$ represents the relative yield function $f_{i}\left(\hat{X}_{i t}^{\top}\right)$, for $i=1,2$. Waggoner and Norvell [1979] fitted an equation depicting the law of the minimum by minimizing the error sum of squares given by

$$
\begin{align*}
\sum_{t}^{N} \varepsilon_{t}^{2}= & \sum_{t}^{n_{1}^{1}\left(y_{t}-f_{1 t}\right)^{2}} & & \text { if } f_{1 t}<f_{2 t} \text { and } A \\
& +\sum_{t}^{n_{2}^{2}\left(y_{t}-f_{2 t}\right)^{2}} & & \text { if } f_{2 t} \leqq f_{1 t} \text { and } A \\
& +\sum_{t}^{n_{3}^{3}}\left(y_{t}-A\right)^{2} & & \text { if } A<f_{1 t} \text { and } f_{2 t} \tag{3.26}
\end{align*}
$$

In equation (3.26), the authors' notation is changed slightly in order to conform to that of equation (3.25). The problem with the approach
of estimation in (3.26) is that an iterative procedure is required to assign subsamples $n_{1}, n_{2}$, and $n_{3}$ to the observations limited by a factor. Even if this objective is achieved, there is no guarantee that the assignment will be unique. Another problem is that the procedure requires large numbers of observations in each subsample.

Equation (3.25) is equivalent to

$$
\begin{equation*}
\operatorname{Max}_{x} y_{t} \tag{3.27}
\end{equation*}
$$

s.t. $y_{t} \leqq A f_{i t} ; i=1,2 ; t=1, N$

Fertilizer recommendations can be derived from this model by setting up an appropriate objective function and optimizing over $X_{i}^{\top}$ subject to the constraints given in (3.27). The constraints are indeed the individual response functions which in Lanzer's study [1978] were estimated independently. This was made possible by use of prior knowledge of the experimental conditions under which the data were collected. The estimation of the response function for a given nutrient was done using only data from the plots in which the other nutrients were known to have been applied at non-limiting levels. This procedure, like the approach adopted by Waggoner and Norvell, is inefficient in terms of data utilization.

Equation (3.25) can be estimated directly by maximum likelihood techniques if a particular probability distribution is assumed for the error term $\varepsilon_{\mathrm{t}}$ and given that the latter are independently and identically distributed (iid). But first, a definition is in order. Assuming that $E\left(\varepsilon_{t}\right)=0, E\left(y_{t}\right)$ is given by $A \operatorname{Min}\left\{f_{1 t}, f_{2 t}\right\}$ which will now be denoted by $\mu_{t} . \frac{1 /}{}$ Equation (3.25) may then be written as
$\underline{1 /}$ This is basically the mean function approach as opposed to a production frontier approach. In the latter case, the residuals may be composite or strictly one-sided and imply estimation procedures different from those anticipated here.

$$
\begin{equation*}
y_{t}=\mu_{t}+\varepsilon_{t} \tag{3.28}
\end{equation*}
$$

$$
\begin{aligned}
\text { where (i) } \mu_{t} & =A f_{1 t}-S_{1 t} \\
\text { (ii) } \mu_{t} & =A f_{2 t}-S_{2 t} \\
\text { (iii) } 0 & =S_{1 t} S_{2 t} \text {; all } t=1, N \\
\text { and (iv) } S_{1 t} & \geqq 0, S_{2 t} \geqq 0 .
\end{aligned}
$$

The symbols $S_{1 t}$ and $S_{2 t}$ represent slack variables, and because they are restricted to be positive, the mean yield, $\mu_{t}$, will be equal to $A f_{i t}$ whenever $\hat{x}_{i t}^{\top}$ is limiting response. For instance, when $\hat{X}_{1 t}^{\top}$ is the limiting nutrient, $S_{1 t}=0$ and $\mu_{t}=A f_{1 t}$.

With the additional assumption that $\varepsilon_{t} \sim N\left(0, \sigma^{2}\right)$, equation (3.28) may be viewed as the standard regression model to which the technique of maximum likelihood estimation (MLE) could be applied in a straightforward manner. The only distinction is that in this case, there are the non-linear and inequality restrictions stated in (i) through (iv) above. These restrictions together with the joint density functions of $\varepsilon=\left(\varepsilon_{1} \cdot \cdot \varepsilon_{N}\right)^{\prime}$ could be used to construct a Lagrangean objective function whose first order conditions could then be used to obtain estimates of the parameters of (3.28). This derivative approach of obtaining MLE leads to awkward objective and constraint functions thus creating numerical difficulties in estimation.

The residuals ( $\varepsilon_{\mathrm{t}}$ ) are assumed to be independently and identically distributed with a common mean and variance. Thus, directly minimizing the error sum of squares, $S(\varepsilon)$, from (3.28) subject to the given restrictions is equivalent to maximizing the log likelihood function of $\varepsilon$ subject to the same restrictions. The procedure will be elaborated shortly.

### 3.4.2 Using linear splines to approximate crop response to individual nutrients

The approach outlined above may be further simplified by linearizing all or some of the constraints in (3.28). The ones targeted for linearization are those stated as $\mu_{t}=A f_{i t}-S_{i t}, i=1,2$. The term $f_{i t}$ is short for $f_{i t}\left(\hat{X}_{i t}^{\top} ; \beta_{i}\right)$ where $\hat{X}_{i t}^{\top}$ is the total supply of the $i^{\text {th }}$ nutrient and $B_{i}$ is a parameter to be estimated. As before, $f_{i t}$ is a relative yield function and increases at a decreasing rate from zero towards its upper limit, 1 . The yield given by $y_{t}$ thus increases correspondingly towards the maximum denoted by $A$. This maximum is achieved when the supply of the $i^{\text {th }}$ nutrient has reached the level denoted by $\hat{X}_{i 3}^{\top}$ in Fig. 3.1. Furthermore, given a nutrient supply in excess of $\hat{X}_{i 4}^{\top}$, yield having plateaued, may be depressed below the level $C$. For most cropnutrient combinations, the point $C$ requires unrealistically large doses of fertilizer and hence the region beyond that point is of no economic interest. The relevant segment is that given by $A B$. It is this segment that will be approximated by linear pieces as explained below. In what follows, the circumflex on $X_{i}^{\top}$ will be dropped but it should be remembered that it refers to an estimate of the total supply of the $i^{\text {th }}$ nutrient, not just the applied amount.

Given fixed points (knots) $X_{i j}^{\top}$ with $x_{i 0}^{\top}<x_{i 1}^{\top}<\ldots .<x_{i k+1}^{\top}$, the functions $S_{j}\left(X_{i}^{\top}\right)=y_{j}$ defined over $\left[X_{i j-1}^{\top}, X_{i j}^{\top}\right]$ for $j=1, k+1$ are called linear splines if each is linear [see for example Poirier, 1973; Suits et al., 1978, and Smith, 1979]. The function $S\left(X_{i}^{\top}\right)$, in Fig. 3.1, over the mesh $\left[X_{i 0}^{\top}<X_{i 1}^{\top}<\ldots .<X_{i k+1}^{\top}\right]$ for any value of $X_{i}^{\top}$ is, therefore, given by

Figure 3.1 Approximating Crop response to a Single limiting nutrient by Linear Splines


$$
\begin{gathered}
S\left(X_{i}^{\top}\right)=\beta_{00}+\beta_{i 0} X_{i}^{\top}+\sum_{j=1}^{k} \beta_{i j}\left(X_{i}^{\top}-X_{i j}^{\top}\right) \text { Dij } \\
\text { where } D_{i j}=1 \text { if }\left(x_{i}^{\top}-x_{i j}^{\top}\right)>0 \\
=0 \text { if }\left(X_{i}^{\top}-x_{i j}^{\top}\right) \leqq 0
\end{gathered}
$$

$$
\text { and } k+1=\text { number of knots }
$$

The coefficient $\beta_{i 0}$ gives the slope over the first segment of $S\left(X_{i}^{\top}\right)$ and $\beta_{i j}$ is the change in slope from one interval to the next as shown for $\beta_{i 3}$ at point $B$ in Fig. 3.1. This implies that the slope over the $j{ }^{\text {th }}$ segment is given by

$$
\begin{equation*}
\beta_{i}=\sum_{\tau=0}^{i} \beta_{i t} \tag{3.30}
\end{equation*}
$$

In view of nutrient essentiality and also due to the fact that $X_{i}^{\top}$ refers to the total supply of the $i^{\text {th }}$ nutrient, the response function $f_{i}$, and hence $S\left(X_{i}^{\top}\right)$, passes through the origin. This means that $\beta_{00}=0$. If $\hat{A}$ is an estimate of the maximum yield, then, expected yield at time $t$ is given by

$$
\begin{align*}
E\left(y_{t}\right) & =\mu_{t}=\hat{A} f_{i t}-S_{i t} \\
& =\beta_{i 0} \hat{A} X_{i t}^{\top}=\sum_{j=1}^{k_{i}} \beta_{i j} \hat{A}\left(X_{i j}^{\top}-X_{i j}^{\top}\right) D_{i j}-S_{i t} \\
& =\sum_{j=0}^{k} \beta_{i j} z_{i j t}-S_{i t} \tag{3.31}
\end{align*}
$$

In equation (3.31), $Z_{i 0 t}=\hat{A} X_{i t}^{\top}$ and $Z_{i j t}=\hat{A}\left(X_{i t}^{\top}-X_{i j}^{\top}\right) D_{i j}$ for $j=1, k_{i}$ and $i=1,2$. The dummy variable $D_{i j}$ is set to 1 if $\hat{A}\left(X_{i t}^{\top}-X_{i j}^{\top}\right)>0$ and is zero otherwise. $X_{i j}^{\top}$ for all $j$ are assumed to be known. They may, however, be variables which must be chosen so as to minimize the error sum of squares. As will be shown later, there are model selection procedures which are invariant to the relative "fits" of the respective models in question. Since use of such selection procedures is intended, the
use of splines with variable knots as a means of perfecting the "fit" of the non-substitution model will not be rigorous.

Let $Z_{i}=\left(Z_{i 0} \cdots Z_{i} k_{i}\right)$ be an $N x k_{i}+1$ matrix of the $X_{i}^{\top}$ vector transformed as indicated in the second term in the second line of (3.31). Also let $\beta_{i}^{\prime}=\left(\beta_{i 0} \cdots \cdot \beta_{i k_{i}}\right)$ be the vector of coefficients, an interpretation of these having been given above. In a compact form, $\mu$, the vector of expected yields, is now given by $z_{i} \beta_{i}$. If $\hat{\mu}$ is the estimate of $\mu$ and $e$, the estimate of $\varepsilon$, the vector of residuals from (3.28), then $S(e)$, the error sum of squares, is given by $e^{\prime} e=(Y-\hat{\mu})^{\prime}(Y-\hat{\mu})$ where $Y$ is the $N \times 1$ vector of yields. The problem of estimating $\beta_{j}$ for i $=1,2$ therefore becomes

$$
\begin{align*}
& \operatorname{Min} \quad \frac{1}{2} \varepsilon^{\prime} \varepsilon \\
& \text { s.t. } S_{1 t} S_{2 t}= \text { all } t=1, N \\
& A_{1} X_{N L}+ A_{2} X_{L}=Y * \\
& S_{i t} \geqq 0 \quad i=1,2 \\
& \varepsilon \text { unrestricted } \tag{3.32}
\end{align*}
$$

where

$$
\begin{aligned}
& A_{1}=\left[\begin{array}{ccc}
0 & 0 & I \\
-I & 0 & 0 \\
0 & -I & 0
\end{array}\right] ; \quad A_{2}=\left[\begin{array}{ccc}
I & 0 & 0 \\
-I & Z_{1} & 0 \\
-I & 0 & Z_{2}
\end{array}\right] \\
& X_{N L}=\left[\begin{array}{c}
S_{1} \\
S_{2} \\
\varepsilon
\end{array}\right] ; \quad X_{L}=\left[\begin{array}{c}
\mu \\
\beta_{1} \\
B_{2}
\end{array}\right] \\
& Y_{*}=\left[\begin{array}{c}
Y \\
0 \\
0
\end{array}\right] ; \quad \beta_{i}^{\prime}=\left(\beta_{i 0} \cdots \beta_{i k_{i}}\right) \\
& \beta_{i 0} \geqq 0 ; \beta_{i j} \leqq 0, j \geqq 1, j=0, k_{i} \text { and } i=1,2
\end{aligned}
$$

The total number of knots in the $i^{\text {th }}$ spline function is given by $k_{i}+1$ and $X_{N L}$ and $X_{L}$ are the sets of nonlinear and linear variables, respectively. The set of restrictions $\beta_{i 0} \geqq 0$ and $\beta_{i j} \leqq 0$ for $j \geqq 1$ ensures concavity of the response functions as assumed in the theory of diminishing increments.

The use of linear splines simplifies the computational task by reducing the degree of nonlinearity in the constraint set. Any reasonable functions, linear or nonlinear, could have been substituted for the spline functions $Z_{i} \beta_{i}$. However, it is usually desirable to limit the number of nonlinear constraints in a problem such as (3.32) above, especially for large sample sizes.

Apart from the computational advantage, approximating $f_{i}\left(X^{\top}, \beta_{i}\right)$ by linear splines makes it unnecessary to be committed to a particular mathematical form for the response function.

A general advantage of the formulation in (3.32) over the other techniques of fitting the law of the minimum (see subsection 3.4.1) is that the whole data set is used to simultaneously estimate crop response to the limiting nutrients. There is no need for dividing observations into subsamples corresponding to a limiting factor or selecting only observations for which the unstudied factor was applied at nonlimiting. levels. Furthermore, the formulation is also amenable to extension so as to account for more than two limiting factors of production.

Preliminary analysis of the Purdue experimental data by the author confirmed Professor Barber's conclusion that there was no significant interaction effects between $P$ and K [Barber, 1958]. No attempt is therefore, made to account for interaction in the model presented in (3.32). The present framework of analysis can be modified rather simply (but
at the expense of computational ease) to account for interaction effects - see for example Poirer [1975].

The problem in (3.32) has a nonlinear objective function and one set of nonlinear constraints, $\mathrm{S}_{\mathrm{it}} \mathrm{S}_{2 t}$ for all t , in addition to linear and inequality constraints. The problem is solved by employing a nonlinear programming algorithm developed by Murtagh and Saunders [1980]. The algorithm is code named MINOS/AUGMENTED and is designed to solve large-scale optimization problems involving sparse linear and nonlinear constraints.

### 3.5 A test of the nutrient non-substitution hypothesis

The need to employ both prior information and statistical criteria in model selection cannot be overemphasized. More often than not, only one of these assets is employed, the other either being downplayed or totally neglected. The use of quadratic and square root polynomials for crop response analysis, for instance, gained popularity largely on grounds of simplicity and ease of computation. They also seem to fit the data on crop production rather well. Hence, they appear attractive whenever measures of relative performance are the only means for discriminating among competing mathematical forms.

The most commonly used measure of relative performance of models are the coefficient of multiple correlation ( $R^{2}$ ) and the mean square error (MSE). The $R^{2}$ is not a powerful tool for selecting the best specification particularly when the competing forms are performing equally well. It may also be an ambiguous statistic when the models are estimated under unequal transformations.

By and large, any selection procedure designed to obtain the best specification presupposes knowledge of the true relationship. Models are then chosen on the basis of how close they are to this true relationship as indicated for example by the $R^{2}$ or MSE. The assumption that the true form of the relationship is known or that it exists under some general conceptual terms is rather strong. But, even if such a relationship existed, choice of mathematical forms to approximate it purely on the basis of their relative fit cannot avoid an exercise of subjectivity whenever there are deadlocks.

Selection procedures employing measures of relative performance as discussed above are generally concerned with the subject of discrimination. The objective is to obtain the best mathematical specification of a given relationship. There are times, however, when the interest is in hypothesis testing rather than discrimination. In the former case, the null hypothesis $\left(H_{0}\right)$ is tested against an alternative one $\left(H_{1}\right) . H_{0}$ is either rejected or not rejected at a prescribed probability level of a type I error (significance level). The decision process is restricted to only two possibilities since the truth of one hypothesis implies falsity of the other. The researcher must, therefore, be willing to be committed to one of the models being tested. Such a commitment implies that (only) one of the models is the true specification in terms of obeying prescribed assumptions. Such an approach is not justified if the true form of the relationship being tested is unknown.

In the regression case, a test of $H_{0}: f_{0}\left(y \mid z, \beta, \varepsilon_{0}\right)$ against an alternative $H_{1}: f_{1}\left(y \mid X, \gamma, \varepsilon_{1}\right)$ can employ the classical $F$-statistic based on the Neyman-Pearson likelihood ratio method if either $X$ is orthogonal tp $Z$ or $M_{x} Z=0$. The symbols $X$ and $Z$ refer to the sets of regressors,
$\beta$ and $\gamma$ are the parameters and $\varepsilon_{0}$ and $\varepsilon_{1}$ are the error terms of the respective models. $M_{x}$ is the principal idempotent matrix under the alternative hypothesis and is given by $I-X\left(X^{\prime} X\right)^{-1} X^{\prime}$. Given that all the classical assumptions of the regression model are obeyed under the respective models, the two hypotheses are said to be nested. This means that one can be obtained as the limiting case of the other.

When the above conditions are not met, as in the case when $X$ is not a subset of $Z$ (or vice versa) or $Z=\log X$, the hypotheses are said to be non-nested and the classical F-test is inappropriate [Cox, 1961; Pesaran, 1974].

In what follows, interest will be in testing the non-substitution specification presented in equation (3.28) against a given alternative. The alternative will be taken as the polynomial formulation. The polynomials chosen for this role are the quadratic and square root specifications.

Polynomials have been used in many studies intended for comparing the performance of different mathematical formulations. This tradition of comparing polynomial approximations with models based on biological principles of plant growth will be upheld here. In the present case, the thesis is that polynomials are used to approximate a biological relationship which is better explained, at least theoretically, in terms of the law of the minimum. The polynomials allow substitution among the essential nutrients $P$ and $K$ contrary to theory of plant nutrition. Equation (3.28), which is a generalization of the law of the minimum, does not allow substitution among the major nutrients. A polynomial portrays a symmetrical surface around a unique maximum yield. It would, therefore, fail to capture sharp bends and plateaus in a crop response surface.

A polynomial fitted to a set of data showing a significant plateau can lead to costly positive biases in fertilizer recommendations to producers.

The problem at hand is to statistically test the non-substitution model (in a format capable of accomodating an extended flat crop response surface as implied in the law of the limit) against a polynomial
formulation. The null and alternative hypotheses are, therefore, given as

$$
\begin{align*}
H_{0}: & f_{0}(Z, \beta)=\mu+\varepsilon_{0} ; \mu=Z_{1} \beta_{1}-S_{1} ; \mu=Z_{2} \beta_{2}-S_{2} ; \\
& S_{1 t} S_{2 t}=0, \text { all } t ; \\
& S_{1} \geqq 0 ; S_{2} \geqq 0 ; \beta_{i 0} \geqq 0 ; \beta_{i j} \leqq 0, j \geqq 1 \\
H_{1}: & f_{1}\left(X^{\top}, \gamma\right)=X^{\top}{ }_{\gamma}+\varepsilon_{1} \tag{3.33}
\end{align*}
$$

The symbols in (3.33) represent variables already defined in the preceding equations and $X^{\top} \gamma_{\gamma} \gamma_{1} X_{1}^{\top}+\gamma_{2}\left(X_{1}^{\top}\right)^{\frac{1}{2}}+\gamma_{3}\left(X_{1}^{\top} X_{2}^{\top}\right)^{\frac{1}{2}}+\gamma_{4}\left(X_{2}^{\top}\right)^{\frac{1}{2}}+\gamma_{5} X_{2}^{\top}$ for the square root formulation. In the case of the quadratic formulation, $x_{\gamma}^{\top}=\gamma_{1} x_{1}^{\top}+\gamma_{2}\left(X_{1}^{\top}\right)^{2}+\gamma_{3} x_{1}^{\top} x_{2}^{\top}+\gamma_{4}\left(X_{2}^{\top}\right)^{2}+\gamma_{5} x_{2}^{\top}$. Due to the concept of nutrient essentiality and because $X_{i}^{\top}(i=1,2)$ are estimates of total nutrient supply, the intercepts are suppressed.

The null and alternative hypotheses in (3.33) are non-nested and hence the classical F-test is not valid. The literature on fertilizer use presents numerous comparisons of polynomials and other mathematical forms, but the author is not aware of any empirical study in which the non-nested nature of hypotheses such as those in (3.33) was statistically recognized. Comparisons have been done in an hoc manner with relative performance as the major criterion.

Studies have shown that simplicity and performance of polynomials can be matched by models whose parameters are readily interpreted in biological terms. Waggoner and Norvel [1979] for example stated that "In
fact, the law of the minimum fits these yields as well or better than any of the empirical functions used by Heady et al., [1955] (Table 1)." Only the $R^{2}$ was used as the criterion for comparison between the law of the minimum and those used by Heady et al. Obviously, the authors' concern was whether their model was better or worse than the empirical alternatives previously suggested. The selection procedure was restricted to the models at hand. There was no room for the possibility that an alternative not considered together with the data could be used to reject all the hypotheses tested.

The statement made earlier that " . . . relationship is better explained . . . in terms of the law of the minimum" is only a claim. Its validity has to be tested empirically against alternative theories and the evidence embodied in the data set. Validity of the non-substitution model is not being claimed for all fertilizer-crop relationships. Room must be left for the possibility that its mathematical form misspecifies either the law of the minimum or the underlying relationship. It is therefore conceivable that a polynomial together with the data can be sufficient to reject it. Conversely, the non-substitution model can be used to reject the polynomials. And in all this, a third alternative is not ruled out, meaning that the reciprocal rejection can lead to the conclusion that the models tested misspecify crop response to fertilizer nutrients. If this is the case, other formulations must be sought. What all this means is that there is really no uniquely maintained hypothesis. Hence, the roles of the two hypotheses in (3.33) must be switched so that a polynomial becomes the maintained hypothesis. Since the hypotheses are non-nested, procedures for testing non-nested hypotheses will be applied.

The problem of testing the hypotheses in (3.33) can be tackled by first constructing a comprehensive model. For the sake of brevity, the following notation will be adopted: $\mathrm{f}_{\mathrm{i}}(\mathrm{i}=0,1)$ are to be viewed as probability density functions (pdf's), $f_{0}=f_{0}(Z, B)$ and $\hat{f}_{0}=f_{0}(Z, \hat{B})$ while $f_{1}=f_{1}\left(X_{1} \gamma\right)$ and $\hat{f}_{1}=f_{1}(x, \hat{\gamma})$.

Linear nesting of models was discussed by Quandt [1974]. The comprehensive model is obtained as

$$
\begin{equation*}
f_{\theta}(y \mid \beta, \gamma)=(1-\theta) f_{0}+\theta f_{1} \tag{3.34}
\end{equation*}
$$

where $y$ is the dependent variable and $\theta$ is the nesting parameter to be tested for zero and unity. In (3.34) parametric identification may fail, but tests on $\theta$ may still determine departures from the null hypothesis in the direction of the alternative or away from it. The identification problem can be circumvented in a number of ways some of which will be mentioned shortly. A major weakness of the comprehensive model shown above is that it may not in itself constitute a viable theory regarding crop response. The two functions $f_{0}$ and $f_{0}$ may be such that the comprehensive model constructed from them does not make sense. Yet the latter acts as a third alternative whenever $\theta$ is significantly different from zero but lies somewhere between zero and one. Another problem that may arise from (3.34) is that of multicollinearity.

An alternative way of overcoming the identification problem is by using prior information. If for instance $\beta$ and $\gamma$ are known, then $\theta$ in (3.34) becomes both a nesting as well as a testing parameter. The null hypothesis is rejected if $\theta$ is significantly greater than zero but less than one, while the alternative hypothesis is rejected when $\theta$ is less than zero. Values of $\theta$ outside $[0,1]$ are interpreted as a movement beyond $H_{1}$ if $\theta>1$ and away from both $H_{0}$ and $H_{1}$ if $\theta<0$ [Fisher and
and McAleer, 1981]. The possibility of ending up with an artificial model such as (3.34) still exists even with prior knowledge of $\beta$ and $\gamma$. Furthermore, such prior information is hardly available. One may therefore have to employ numerical techniques of identification mentioned below.

An estimate of $\theta$ may be obtained by estimating a comprehensive model in which $\hat{\beta}$ and $\hat{\gamma}$ are substituted for $\beta$ and $\gamma$ giving $y=(1-\theta) \hat{f}_{0}+\theta f_{1}$ $+\varepsilon$. Getting back to the notation of (3.33) this would imply estimating

$$
\begin{equation*}
y-\hat{\mu}=\theta\left(X \mathbf{T}_{\gamma}-\hat{\mu}\right)+\varepsilon \tag{3.35}
\end{equation*}
$$

$H_{0}$ is then tested on the basis of the t-ratio statistic for $\theta$, not on the basis of its absolute value. The estimate of $\theta$ is conditional on those of $\beta$ and $\gamma$, and hence Davidson and MacKinnon [1981] referred to the test based on (3.35) as the C-test. The C-test is a simple way of testing the non-substitution hypothesis. Davidson and MacKinnon stated that the $t$-statistic from (3.35) provides a test the asymptotic size of which is smaller than its nominal size. The C-test, therefore, has a higher probability of type I error. It follows that if $H_{0}$ is rejected by $H_{1}$, this would be strong evidence against the hypothesis under test.

Another test proposed by the same authors is referred to as the $J$ test. It is based on the asymptotic $t$ statistic of $\theta$ estimated from the following comprehensive model

$$
\begin{equation*}
y=(1-\theta) \mu+\theta X^{T_{\gamma}}+\varepsilon \tag{3.36}
\end{equation*}
$$

Unlike equation (3.35) in which both $\hat{B}$ and $\hat{\gamma}$ are used, (3.36) substitutes only for one of the parameters, $\gamma$. The other parameters $\beta$ and $\theta$ must be estimated jointly hence the name J-test. If $H_{0}$ is true, the t-statistic will be asymptotically distributed as $N(0,1)$. A valid statistic for testing the truth of $\mathrm{H}_{1}$ is obtained by reversing the roles of $\mathrm{H}_{0}$ and $H_{1}$ in (3.33).

The J-test will not be employed in this study due to the following reasons. First, equation (3.36) requires estimating $\mu$ and $\theta$ jointly. Since $\mu$, the $E(y)$ in the non-substitution model, is given by $Z_{1} \beta_{1}-S_{1}$ and $Z_{2} B_{2}-S_{2}$ subject to $S_{1 t} S_{2 t}=0$ and non-negative restrictions on $S_{1}$ and $S_{2}$, (3.36) must be estimated using the mathematical programming framework given in (3.32). The programming approach does not provide standard errors for the estimated coefficients. ${ }^{1 /}$ Employing the J-test would, therefore, require lengthy auxiliary estimations which are avoidable under alternative testing procedures. The second reason for not employing the J-test is that when $H_{0}$ is true, the statistic has a perfect negative asymptotic correlation with the Cox-Pesaran (CP) statistic (to be discussed shortly). The J-statistic is, therefore, basically the same as the CP statistic in terms of inferences [Davidson and MacKinnon, 1981]. 2/ The CP procedure is easily adaptable to the mathematical programming framework in (3.32).

There are other alternative procedures for testing non-nested hypotheses. Such procedures will not be discussed here in detail. A good summary can be found in Gaver and Geisel [1974] who also present Bayesian techniques.

Exponential nesting of models was first proposed by Cox [1961] and employed by Atkinson [1970] for discriminating among alternative

1/Given that $\varepsilon t$ is iid and that the sample size is reasonably large, the desired standard errors may be obtained from the information matrix, $-E\left[\partial^{2} L(\theta) / \partial \theta \partial \theta^{\prime}\right]$, constructed from equation (3.28) and the accompanying restrictions. $L(\theta)$ is the log likelihood function and $\theta$ is the parameter vector.

2/ The statement is valid only when $H_{0}$ is true and for large samples. For small sample sizes the two procedures are different and may yield conflicting results.
specifications. It has also been discussed by Quandt [1974] and more recently by Fisher and McAleer [1981]. In exponential nesting, the comprehensive model is given by $f_{\theta}\left(y \mid \beta_{1} \gamma\right)=k f_{0}^{1-\theta} f_{1}^{\theta}$ where for $f_{\theta}(\cdot)$ to be a true pdf, the following restriction must hold:

$$
1 / k=\int_{-\infty}^{\infty} f_{0}^{1-\theta} f_{1}^{\theta} d y
$$

Under the appropriate (numerical) identifying restrictions, the J-statistic may be derived from an artificial model obtained by nesting models exponentially. As already noted, the J-statistic is not adaptable to the techniques of estimation in this study.

The procedure to be discussed next was first proposed by Cox [1961] and has since been elaborated by Pesaran [1974] for linear regression models and Pesaran and Deaton [1978] for nonlinear regression models. The procedure will simply be referred to as $C P$. The test statistic derived from the $C P$ procedure is given by

$$
\begin{equation*}
\left.T_{0}=\hat{L}_{01}-N[P] \lim _{N \rightarrow \infty}\left(\hat{L}_{01} / N\right)\right]_{\beta=\hat{\beta}} \tag{3.37}
\end{equation*}
$$

which is shown by Cox [1961] to be asymptotically normally distributed with zero mean and variance $V_{0}\left(T_{0}\right)$, given that the null hypothesis is true. $\hat{L}_{01}=L_{0}(\hat{\beta})-L_{1}(\hat{\gamma})$ where $L_{0}(\hat{\beta})$ and $L_{1}(\hat{\gamma})$ are the $\log$ likelihood functions of samples of size $N$ under the null and alternative hypotheses. $\hat{\beta}$ and $\hat{\gamma}$ are MLE of $\beta$ and $\gamma$ respectively. It follows that

$$
\begin{equation*}
D_{0}=T_{0} /\left[V_{0}\left(T_{0}\right)\right]^{\frac{1}{2}} \tag{3.38}
\end{equation*}
$$

is a standardized normal variate that can be used in conjunction with regular statistical tables of the normal distribution to test the truth of $\mathrm{H}_{0}$.

Pesaran [1974] showed that the statistic in (3.37) for linear regression models is given by

$$
\begin{align*}
T_{0} & =\frac{N}{2} \log \left(\hat{\sigma}_{1}^{2} / \hat{\sigma}_{10}^{2}\right) \\
& =\frac{N}{2} \log \left[\hat{\sigma}_{1}^{2} /\left(\hat{\sigma}_{0}^{2}+\frac{1}{N} e_{10}{ }^{\prime} e_{10}\right)\right] \tag{3.39}
\end{align*}
$$

where $-\sigma_{10}^{2}$ is the asymptotic expectation of $\hat{\sigma}_{1}^{2}$ under $H_{0}$. The respective sample variances $\hat{\sigma}_{i}^{2}$ for $i=0,1$ are given by $e_{i}^{\prime} e_{i} / N, e_{i}$ being the residual vector of the $i^{\text {th }}$ hypothesis. In the case of the hypothesis testing problem in (3.33), the vector $e_{10}$ is obtained from the following regression:

$$
\begin{equation*}
\hat{\mu}=X^{\top} b_{0}+\varepsilon_{10} \tag{3.40}
\end{equation*}
$$

where $\mu$, the estimate of $E(y)=\mu$ for the non-substitution model, is obtained from (3.32) and $X^{\top}$ is the set of regressors for the polynomial model.

The variance of $T_{0}$ is given by

$$
\begin{align*}
V_{0}\left(T_{0}\right) & =\frac{\hat{\sigma}_{0}^{2} e_{100}{ }^{\prime} e_{100}}{\hat{\sigma}_{10}^{4}} \\
& =\frac{\hat{\sigma}_{0}^{2} e_{100}{ }^{\prime} e_{100}}{\left(\hat{\sigma}_{0}^{2}+\frac{1}{N} e_{10}{ }^{\prime} e_{10}\right)^{2}} \tag{3.41}
\end{align*}
$$

where the vector $e_{100}$ is obtained from the following regression:

$$
\begin{equation*}
e_{10}=\mu+\varepsilon_{100} \tag{3.42}
\end{equation*}
$$

To test the truth of $H_{0}$ using $D_{0}$ in (3.38), one would therefore require the auxiliary regressions in (3.40) and (3.42). The regression in (3.40) simply requires substituting $\hat{\mu}$ on the left hand side of the program package used for estimating the polynomial model. This regression yields ${ }^{\mathrm{e}}{ }_{10}$ which is then used as the right hand side column vector in the MINOS/ AUGMENTED program employed in estimating (3.32). The final estimate of $D_{0}$ in (3.38) is obtained by substituting $\bar{\sigma}_{0}^{2}=e_{0}{ }^{\prime} e_{0} /\left(N-K_{0}\right)$ and
$\bar{\sigma}_{1}^{2}=e_{1}^{\prime} e_{1} /\left(N-K_{1}\right)$ for $\hat{\sigma}_{0}^{2}$ and $\hat{\sigma}_{1}^{2}$, where $N-K_{0}$ and $N-K_{1}$ are the respective degrees of freedom under $H_{0}$ and $H_{1}$. This final value of $D_{0}$ will be referred to as $\bar{D}_{0}$ to distinguish it from the one defined in (3.38)

Since the parameter spaces for the two models being tested are disjointed, a two tailed test will be employed. For a given level of significance, $\alpha$, if the tabulated value of the statistic is given by $D_{\alpha}$, then $H_{0}$ is not rejected if $\left|\tilde{D}_{0}\right|<\left|D_{\alpha}\right| . H_{0}$ is rejected in favor of $H_{1}$ if $\left|\bar{D}_{0}\right|>\left|D_{\alpha}\right|$ and $\bar{D}_{0}$ is negative. Finally, if $\left|\bar{D}_{0}\right|>\left|D_{\alpha}\right|$ and $\bar{D}_{0}$ is positive, the null hypothesis is rejected but in favor of some alternative $H_{2}$ differing from $H_{0}$ in some sense opposite to that in which $H_{1}$ differs from $H_{0}$.

The statistic $\overline{\mathrm{D}}_{0}$ is only valid for testing the truth of $\mathrm{H}_{0}$. In order to test the truth of $H_{1}$ the roles of the hypotheses are reversed in (3.33). A new statistic is then computed estimating two more auxiliary regressions. The new statistic will be referred to as $\overline{\mathrm{D}}_{1}$. The regression which corresponds to (3.40) is now given by

$$
\begin{equation*}
x^{\top} \hat{\gamma}=\mu+\varepsilon_{01} \tag{3.43}
\end{equation*}
$$

Equation (3.43) provides the residuals $e_{01}$. It requires replacing the right hand side of (3.32) by $X^{\top}{ }_{\gamma}$. Next, the residuals are substituted on the left hand side of the program package used for estimating the polynomial model to obtain the equation

$$
\begin{equation*}
e_{01}=x^{T} b_{1}+\varepsilon_{011} \tag{3.44}
\end{equation*}
$$

Equation (3.44) provides the residuals $\mathrm{e}_{011}$. The new statistic which tests the truth of the polynomial specification is then computed as originally done for $\bar{D}_{0}$.

Equation (3.37) can be written alternatively as

$$
\begin{equation*}
T_{0} \equiv T_{0}(\hat{\Gamma})=\left(\hat{L}_{0}-\hat{L}_{1}\right)-E_{0}\left(\hat{L}_{0}-\hat{L}_{1}\right) \tag{3.45}
\end{equation*}
$$

where $\hat{\Gamma}=\left(\hat{\gamma} \hat{\sigma}_{1}^{2}\right)$. This leads to the relationship in (3.39). $\hat{L}_{0} \equiv L_{0}(\hat{\beta})=N / 2 \log \left(2 \pi \hat{\sigma}_{0}^{2}\right)-N / 2$ is independent of $\Gamma$. It does not matter, therefore, what value is assigned to $\hat{F}$ provided it is consistent for $\Gamma .-$ In (3.45) MLE $\hat{\Gamma}$ is used. One may, however, replace $\Gamma$ with a consistent estimate of $\Gamma_{0}$, the asymptotic expectation of $\hat{\Gamma}$ under $H_{0}$. This leads to a different numerator for $D_{0}$ and is given by

$$
\begin{equation*}
T A_{0}=T_{0}+\left(\hat{L}_{1}-\hat{L}_{10}\right) \tag{3.46}
\end{equation*}
$$

Atkinson [1970] shows that $T A_{0}$ and $T_{0}$ are asymptotically equivalent under $H_{0}$. Fisher and McAleer [1981] derived $T A_{0}$ as

$$
\begin{equation*}
T A_{0}=\frac{N}{2}\left\{\left(\hat{\sigma}_{1}^{2} / \hat{\sigma}_{10}^{2}\right)-1\right\}+\frac{1}{2} \hat{\sigma}_{10}^{2}\left\{\left(y-\hat{f}_{10}\right)^{\prime}\left(y-\hat{f}_{10}\right)-e_{1}^{\prime} e_{1}\right\} \tag{3.47}
\end{equation*}
$$

where $\hat{f}_{10}=f_{1}\left(\hat{\gamma}_{0}, \hat{\sigma}_{10}^{2}\right)$ and $\gamma_{0}$ and $\sigma_{10}^{2}$ are the asymptotic expectations of $\hat{\gamma}$ and $\hat{\sigma}_{1}^{2}$ under $H_{0}$, respectively.

Since the expression of $T_{0}$ in (3.39) is approximated by

$$
\begin{align*}
T L_{0} & =\frac{N}{2}\left\{\left(\hat{\sigma}_{1}^{2} / \hat{\sigma}_{10}^{2}\right)-1\right\} \\
& \geqq \frac{N}{2} \log \left(\hat{\sigma}_{1}^{2} / \hat{\sigma}_{10}^{2}\right) \tag{3.48}
\end{align*}
$$

the following relationships must hold:

$$
\begin{equation*}
T A_{0} \geqq T L_{0} \geqq T_{0} \tag{3.49}
\end{equation*}
$$

All the three variations of the Cox-Pesaran statistic are asymptotically equivalent under $H_{0}$, and if a common variance $V_{0}\left(T_{0}\right)$ in (3.41) is assumed, the relationship in (3.49) implies that

$$
\begin{equation*}
D A_{0} \geqq D L_{0} \geqq D_{0} \tag{3.50}
\end{equation*}
$$

Fisher and McAleer [1981] concluded that "When the alternative, $H_{1}$, is fitting much better (worse) than it ought, relying solely on $D_{0}\left(D A_{0}\right)$ will more likely lead to rejection of $H_{0}$ than would otherwise be the case." The linearized statistic $\mathrm{DL}_{0}$ is more conservative at rejecting $H_{0}$
than is $D_{0}\left(D A_{0}\right)$ when $H_{1}$ is fitting much better (worse) than might be expected.

Since $L_{0}$ requires only a slight modification of the numerator for $D_{0}$, it will also be computed. Thus, to recap, the techniques to be employed are the C-test as a preliminary test of $\mathrm{H}_{0}$ and the two variations ( $\mathrm{D}_{0}$ and $D L_{0}$ ) of the CP procedure. It is also worth noting that CP procedures involve computing the test statistic under only one of the hypotheses, and hence there is no choice involving an artificial (comprehensive) model. The test, however, indicates whether there is a more appropriate specification beyond $H_{1}$ or away from both $H_{0}$ and $H_{1}$. The test may therefore reject both hypotheses while indicating a direction in which to search for alternatives.

## Chapter 4: Empirical Results

### 4.1 Fertilizer carryover functions

### 4.1.1 The phosphorus carryover function

The equation estimated was of the form

$$
\begin{align*}
S P_{j}= & b_{1}+b_{2}+b_{3}+b_{4}+(\rho+\gamma) S P_{j-1}-\rho \gamma S P_{j-2} \\
& +(1-\rho) \gamma \lambda_{4}^{-2} C P_{j-1}+(1-\rho) \gamma \lambda_{2}^{-1} W P_{j}+\alpha_{1} \text { HAY }_{j} \\
& -\rho \alpha_{1} \text { HAY }_{j-1}+\alpha_{2} \text { WHEAT }_{j}-\rho \alpha_{2} \text { WHEAT }_{j-1} \\
& +\alpha_{4} \text { CORN }_{j}-\rho \alpha_{4} \text { CORN }_{j-1}+W_{j} \tag{4.1}
\end{align*}
$$

where $j$ refers to a rotation and the dependent variable SP is the soil test phosphorus measured in $\mu \mathrm{g} / \mathrm{g}$ by Bray P1. The error term is assumed to be white noise with a distribution $W_{j} \sim N\left(0, \sigma^{2} I\right)$ for all $j$.

Soil sampling was not done for the years 1970, 1971, and 1972. Estimation of both $P$ and $K$ carryover functions was, therefore, restricted to the period 1953-1969. The period consists of 4 complete rotations and one quarter of the fifth rotation (1969). During this period of experimentation, the application rates for $P$ in wheat (WP) and corn-1 $(C P)$ were constant. $W P_{j}$ and $W P_{j-1}$ are therefore identical and so are $C P_{j-1}$ and $C P_{j-2}$. The term ( $1-\rho$ ) in equation (4.1) arises from collection of like terms on the basis of this fact as can be deduced from equation (3.24). Preliminary computer runs indicated that results were not significantly different if, instead, $W P_{j-1}$ and $C P_{j-2}$ were omitted.

The block dummy variables denoted by $b_{i}(i=1,4)$ were included in equation (4.1) to account for the influences of the different blocks. Such differences may be due to block to block variations in seasonal
availability of $P$. The inclusion of the dummy variables (all the four since there was no natural constant in the equation) led to a significant improvement in the fit as judged from the residual sum of squares (SSR). ${ }^{1 /}$

Except for the soil test, the values for all the other variables were transformed to $\mathrm{kg} / \mathrm{ha}$ as discussed in the appendix A.l. Estimation of equation (4.1) was done by use of SHAZAM. The results are given in Table 4.1.

The provision of reasonable starting values for the coefficients facilitated rapid convergence after 23 iterations and 35 function evaluations. The starting values were obtained from the results of estimating a linearly restricted version of equation (4.1). The results for the estimation of the restricted equation are not presented here because the derived coefficient estimates are not unique.

The primary purpose for estimating the carryover function in (4.1) is to obtain a consistent estimate of the parameter $\lambda$. This is the proportionality constant required in the calibration of soil test values. Subsequent estimations leading to the test of the non-substitution hypothesis will be concentrating on the response of corn-1 to total nutrient supply. The objective here is to derive the relevant value of $\lambda$ to be used for constructing total P. The total P, PT, available for the corn-1 crop is given by $\hat{\lambda}_{p} S P+C P$ where the soil test variable $S P$ is in $\mu \mathrm{g} / \mathrm{g}$ and $C P$ is in $\mathrm{kg} / \mathrm{ha}$ of applied $P$.

The basic time framework in equation (4.1) is the rotation which
$\underline{1 /}$ The null hypothesis that $S_{S R}=\operatorname{SSR}$ was rejected on the basis of an $F$-statistic given by $F=\left[\left(S S R-S S R_{D}\right) / K_{D}-K\right] \div S S R_{D} \mid N-K D$ where the values from the equation with the dummy variables are subscripted by $D$. K refers to the number of coefficients and $N=352$ is the sample size. The estimated statistic was 6.4 which, at $5 \%$ significance level, is greater than 2.37, the tabulated value of $F$ with 4 and 340 degrees of freedom.
consisted of four growing seasons. The coefficient estimates in Table 4.1 therefore refer to four seasons. The response function for corn-1 will, however, be estimated using a single growing season as the basic time framework. It follows that the relevant $\lambda$-value for $P$ must be extrapolated from those given in Table 4.1.

The estimate of the proportionality constant $\lambda_{4}$ given in Table 4.1 is relevant for $P$ applied four seasons back while the value for $\lambda_{2}$ refers to $P$ applied two seasons prior to soil sampling. The extrapolation of a single season $\lambda$-value denoted by $\hat{\lambda}_{p}$ from $\hat{\lambda}_{4}=33.11$ and $\hat{\lambda}_{2}=20.16$ is illustrated in Fig. 4.1.

As already explained in an earlier section, the proportionality constant $\lambda_{p}$ can be interpreted as the proportion of soil nutrient to be made available to the crop during the succeeding season. Only part of the applied $P$ is taken up by the crop. The other part is retained in the soil and released to subsequent crops at a rate dependent on such factors as soil and weather conditions, management, crop characteristics and length of the growing season. One may expect that the longer the growing period the larger is the proportion of fixed $P$ made available to the crop and hence the larger is the proportion $\lambda_{p}$, ceteris paribus. In this sense, it is to be expected that $\lambda_{4} \geqq \lambda_{2}$. The field operations including fertilization and soil sampling were routinely carried out at approximately the same time within each period of a rotation. It was therefore hypothesized that the mean value of $\lambda_{2}$ would be about half that of $\lambda_{4}$ and that the desired single season $\lambda$-value would be about half that of $\lambda_{2}$. The value of $\lambda_{\mathrm{p}}^{*}=P C^{*}$ in Fig. 4.1 would therefore have to satisfy the relationship $Q^{\prime *} / R A=P C * / Q B^{*}$ or $P C^{*}=Q B^{*} / R A Q B^{*}$ and since it is assumed that $Q B^{\star}=\frac{1}{2} R A$, this implies that $\lambda_{p}^{\star}=\frac{1}{2} Q B^{\star}=\frac{1}{2}(16.57)$.

Table 4.1 Results for the Estimation of the Phosphorus Carryover Function (Sample Size $N=352$ )

Coeff.
Coefficient estimates (t-ratios)

|  |  | (a) | (b) | (c) |
| :---: | :---: | :---: | :---: | :---: |
| BLOCK DUMMY | $b_{1}$ | -- | 5.8509(1.9) | 5.7427 (2.0) |
|  | $\mathrm{b}_{2}$ | -- | $3.2407(1.0)$ | 3.2865(1.1) |
|  | $b_{3}$ | -- | 6.6622(2.1) | $6.6034(2.2)$ |
|  | $b_{4}$ | -- | 5.2869(1.8) | $5.3739(1.9)$ |
| AUTOREGRSVE | $\rho$ | -0.1366(2.1) | -0.1192(1.9) | -0.1356(2.0) |
| DISTRIBUTED LAG | $\gamma$ | $0.6954(10.8)$ | ) $0.6776(8.6)$ | 0.6898(10.4) |
| PROPORT. CONST. | $\lambda_{4}$ | 30.5810(5.2) | 31.7460(3.8) | 33.1389 (4.7) |
|  | $\lambda_{2}$ | 19.9203(3.7) | 19.1939(3.1) | 20.1593(3.7) |
| HAY/CORN-2 | $\alpha_{1}$ | -0.0008(5.1) | -0.0011(5.9) | -0.0011(7.4) |
| WHEAT | $\alpha_{2}$ | 0.0010(3.0) | $0.0001(0.2)$ | $0.0002(0.5)$ |
| SOYBEAN | $\alpha_{3}$ | -0.0004 (0.5) | 0.0004(0.5) | -- |
| CORN-1 | $\alpha_{4}$ | 0.0005(3.2) | 0.0005(2.2) | $0.0005(2.5)$ |
| EXTRAPOLATED | ${ }^{\prime} \mathrm{K}$ | 12.98 | 11.61 | 12.26 |
| LOG LIKELIHOOD FN. |  | -1021.89 - | -1013.05 | -1013.25 |
| RVE* |  | 19.91 | 18.51 | 18.53 |

*RVE is the residual variance estimate obtained as SSR/N-K where $N=352$ is the sample size, $S S R$ is the residual sum of squares and $K$ is the number of coefficients estimated.

Figure 4.1 - Extrapolation of a Single Season Proportionality Constant for Phosphorus ( $\lambda_{p}$ )


The estimated value of $\lambda_{p 2}$ is slightly higher than half that of $\lambda_{p 4}$, the former being represented by $Q B$ in Fig. 4.1. The point $B$ is located on line $A D$ where $O D$ represents the proportion of $P$ available at the beginning of the season. The extrapolated proportion of $P$ at the start of the season would be zero if $\lambda_{p}$ had to be derived as discussed in the above paragraph, that is, if the line $A 0$ is adopted. Such a conclusion would not be too realistic.

The advantage of treating differently the phosphorus applied at different periods of a rotation, as it turns out now, is that it helps in locating an extra point $(B)$. Without this point, a linear extrapolation of $\lambda_{p}$, purely on the basis of $\lambda_{p 4}$, may lead to an underestimation of the proportion of soil $P$ made available to a crop during a single growing season. The correct value of $\lambda_{p}=P C$ in Fig. 4.1 must, therefore, satisfy the relationship $Q B / R A=P C / Q B$ or $P C=Q B / R A \cdot Q B$. Since $Q B=20.17$ and $R A=33.14$, the desired value for $\lambda_{p}$ is 12.26 .

Given the above value for $\lambda_{p}$ and a soil test value $S P$, the phosphorus carried over one season to the next (PC) is obtained as $P C=$ $\lambda_{p} S P=12.26 S P$. By definition, therefore, $\lambda_{p}$ is equal to the ratio $P C / S P$ which is in $\mathrm{kg} /$ ha of applied $P$ for every unit of $\mu \mathrm{g} / \mathrm{g}$ of the soil test value. That is, $\lambda_{p}$ is the amount of applied $P$ required to raise the soil test level by one unit.

Barber [1979] reported, for the same set of data used here, that Bray P1 increased by one $\mu \mathrm{g} / \mathrm{g}$ for every $17.0 \mathrm{~kg} / \mathrm{ha}$ of P added. He obtained this figure as the reciprocal of a slope coefficient of the regression of Bray $P 1$ on calculated net change in $P$ over a period of 25 years. The computation of the net change in $P$ required auxiliary laboratory analysis of plant samples which is not necessary if the approach
of the carryover function is to be used. There are many differences, both statistical and technical, between Barber's approach and that used in this study, hence it is not surprising that the two results differ. The estimated $\lambda$-values are, however, of the same magnitude.

As a means of judging the consistency of the results of estimation of equation (4.1) it is worth comparing the extrapolated $\lambda$-value with that obtained by Lanzer [1978] for the south Brazilian data. Lanzer used a carryover function similar to the one used in this study. He reported a $\lambda$-value of 48.26 .

Unlike the Indiana soil which was relatively saturated by $P$ due to repeated application of $\mathrm{P}_{2} \mathrm{O}_{5}$, the Brazilian soils showed lower soil test values for $P$. These soils require more of applied $P$ per hectare to raise the soil level of $P$ by an amount equivalent to one part per million of Bray P1.

The advantage of the carryover function over other techniques of estimating total nutrient supply is that it provides useful additional information. Table 4.1, for example, has estimates of the autoregressive and distributed lag parameters $\rho$ and $\gamma$. The reported values are on the basis of four period rotations. Single period estimates of the absolute values of these parameters are obtained as $\left|\hat{\rho}_{p}\right|={ }^{4} v{ }_{v} \hat{\rho}_{p} \mid=0.6068$ and $\left|\hat{\gamma}_{p}\right|=4 \sqrt{\left|\gamma_{p}\right|}=0.9113$.

The autoregressive coefficient $\hat{\rho}_{p}$ is negative implying a negative influence of a given season on the contiguous seasons. The estimate of the distributed lag parameter is positive and has an absolute value less than one as required. The results show that for the Indiana soil investigated here, only about $10 \%$ of the applied P is taken up by the crop in a single season, the remainder being carried over. This explains the
high soil test values recorded in the plots which were receiving high doses of applied P. In such plots, a Bray Pl of $30 \mu \mathrm{~g} / \mathrm{g}$ was typical and given that $\lambda_{p}=12.26$, a plot receiving $470 \mathrm{~kg} / \mathrm{ha}$ of broadcast $P$ would have approximately $840 \mathrm{~kg} /$ ha of seasonal supply of total P. This by any standard, is substantially beyond the amount required for maximum response of a corn crop.

Finally, results for the estimation of the carryover function show the influence of the various crops on soil phosphorus buildup, information which may prove to be useful in designing a rotation sequence.

Soybean did not seem to have any significant effect on the carryover of $P$ from one growing season to the other. Wheat had an unexpected significant positive coefficient as can be seen from column (a) of Table (4.1). This result seems, however, to have been spurious from a statistical point of view because wheat had no significant influence on the soil test level of $P$ after the block effects were isolated - see columns (b) and (c) of Table 4.1 and the corresponding residual variance estimates (RVE). In spite of the insignificant coefficient estimates for both wheat and soybean in the complete model (column (b)), the results used in this study are derived from a model in which only soybean is omitted (column (d)). The reason is that the omission of soybean, unlike that of wheat, had no destabilizing effect on the estimates of the included coefficients.

Hay/corn-2, the crop harvested just prior to the soil sampling period, had a significant negative effect on phosphorus carryover; the higher the yield the lower is the succeeding soil test value.

Corn-1 planted four seasons prior to soil sampling had a significant positive effect on the soil test value. In an earlier section, it was postulated that removal of crop material from the experimental plots
would have a negative effect on the succeeding soil test value. This statement may be true only if the harvested portion of the crop constitutes a significant loss of $P$. A positive net loss in measurable soil P resulting from the harvest of the hay/corn-2 crop just before soil sampling may explain why $\hat{\alpha}_{1}$ has a negative sign. But it is not so easy to explain why $\alpha_{4}$, the estimated coefficient for corn-1, is positive. One possibility is that the decomposed trash may have boosted the soil level of $P$. If this line of reasoning is acceptable then it follows that the full benefits of trash are not realized inside of two growing seasons as judged from the fact that wheat showed no significant effect on the buildup of soil P. Soybean which is harvested three periods prior to soil sampling has a low trash to harvested crop yield ratio relative to the other crops in the rotation sequence. This may explain why it had no significant effect on the phosphorus carryover.

All the three variations of the phosphorus carryover functions discussed here were fairly stable both in terms of convergence (given different starting values for the coefficients) and the extrapolated value of $\lambda_{p}$.

For a given block ( $B$ ), the results in Table 4.1 column ( $C$ ) can be summarized by the help of equation (4.1) as follows:

$$
\begin{align*}
S P_{j}= & B+0.5542 S P_{j-1}+0.0935 S P_{j-2}+0.0236 C P_{j-1} \\
& +0.0389 W P_{j}-0.0011(\text { HAY } / C O R N-2)_{j} \\
& +0.0002(\text { HAY /CORN-2) } \\
& +0.0001(\text { CORN-1 })_{j-1}+W_{j} \tag{4.2}
\end{align*}
$$

Where $B$ is a dummy variable to be substituted by the relevant coefficient obtained from column (c) of Table 4.1. An alternative way of presenting the same information is by use of the reduced form equation given in
(3.17), subsection 3.3.2. The error term in that equation is then given by

$$
\begin{align*}
V_{j} & =U_{j}-\gamma U_{j-1} \\
& =U_{j}-0.6898 U_{j-1} \tag{4.3}
\end{align*}
$$

where $U_{j}=\rho U_{j-1}+W_{j}=-0.1356 U_{j-1}+W_{j}$.

### 4.1.2 The potassium carryover function

In estimating the phosphorus carryover function, equation (4.1), it was assumed that $W_{j}$ was white noise. The rationale behind this was that the equation was derived on the basis of the assumption that the distributed lag parameter ( $\gamma$ ) and the autoregressive parameter ( $\rho$ ) were not equal. There was no evident indication for the need to simplify the estimated equation on the basis of an assumption that $\gamma=\rho$. In the case of potassium, there was no clear-cut distinction between these two assumptions.

The residual variance estimates given in columns (a) and (b) of Table 4.2 show that there was a definite advantage of isolating the block effects. The results in column (b) were obtained from an equation identical to that given in (4.1) after substituting SK, CK, and WK for SP, CP, and WP respectively. Column (a) summarizes results for the estimation of the same equation without the block dummy variables.

Apart from the negative influence on the residual variance estimate, the exclusion of the dummy variables had some noteworthy effects. The first of these is the switch in significance roles between $\gamma$ and $\rho$. In the complete model, column (b), $\gamma$ is only marginally significant at the 2.5\% level while $\rho$ is highly significant. When the dummy variables are

Table 4.2 Results for the Estimation of the Potassium Carryover Function ( $\mathrm{N}=352$ )

Coeff. Coefficient estimates (t-ratios)

|  |  | $(a) \rho \neq \gamma$ | $(b) p \neq \gamma$ | $(c) \rho=\gamma$ |
| :--- | :--- | :---: | :---: | :---: |
| BLOCK DUMMY | $b_{1}$ | -- | $44.0280(6.7)$ | $60.1460(7.6)$ |
|  | $b_{2}$ | -- | $31.0040(4.5)$ | $44.5680(5.2)$ |
|  | $b_{3}$ | - | $43.7190(6.7)$ | $60.2060(7.2)$ |
|  | $b_{4}$ | -- | $49.6960(7.7)$ | $64.4970(8.0)$ |
| AUTOREGRSVE | $\rho$ | $0.0795(.65)$ | $0.4100(7.8)$ | - |
| DISTRIBUTED LAG | $\gamma$ | $0.3874(3.2)$ | $0.0214(2.1)$ | $0.3513(5.3)$ |
| PROPORT. CONST. | $\lambda_{4}$ | $8.4175(2.7)$ | $0.4521(2.2)$ | $6.3635(4.1)$ |
| HAY/CORN-2 | $\lambda_{2}$ | $4.3937(1.8)$ | $0.2288(1.9)$ | $3.1898(2.8)$ |
| WHEAT | $a_{1}$ | $0.0001(.1)$ | $-0.0027(7.9)$ | $-0.0030(5.5)$ |
| SOYBEAN | $a_{2}$ | $0.0059(3.4)$ | $-0.0012(0.9)$ | $-0.0002(0.1)$ |
| CORN-1 | $a_{3}$ | $-0.0080(3.8)$ | $-0.0050(2.8)$ | $-0.0061(2.8)$ |
| EXTRAPOLATED | $\lambda_{4}$ | $0.0016(3.0)$ | $-0.0006(1.3)$ | $-0.0001(0.2)$ |
| LOG LIKELIHOOD FN. |  | 2.29 | 0.12 | 1.60 |
| RVE |  | 234.81 | 135.05 | 143.36 |

excluded, $\rho$ becomes insignificant and $\gamma$ significant. In either case, the significant coefficient is approximately equal to 0.4 . The second effect to be observed in column (b) is that all the crops indicated a negative influence on the soil test value of $K$ even though only hay/corn-2 and soybean had significant coefficients at the $2.5 \%$ level. When the dummy variables were excluded, however, only soybean showed a negative influence on the soil test value of $K$.

The set of results shown in column (a) in which most of the crops have a positive or zero effect on the carryover of $K$ is unusual. It is improbable, for instance, that the hay/corn-2 crop which is harvested just prior to the soil sampling period could have an insignificant effect on the soil test value. The sensitivity of the coefficient estimates to the exclusion or inclusion of the block dummy variables, the unexpected signs on three out of four of the crop coefficients and the fact that both $\gamma$ and $\rho$ were only significant at a value close to 0.4 suggested a need for a respecification of the potassium carryover function.

The contribution of the block dummy variables toward the improvement of the fit could not be disputed. But, because of the above mentioned change in roles or shift in significance between $\gamma$ and $\hat{\rho}$, it became necessary to investigate the influence of the assumption that $\gamma=\rho$. Under this assumption, the equation estimated was

$$
\begin{align*}
S K_{j}= & b_{1}+b_{2}+b_{3}+b_{4}+\gamma S K_{j-1}+\gamma \lambda_{4}^{-1} C K_{j-1}+\gamma \lambda_{2}^{-1} W K_{j} \\
& +a_{1} H A Y_{j}+a_{2} \text { WHEAT }_{j}+\alpha_{3} S O Y_{j}+\alpha_{4} C O R N_{j}+V_{j} \tag{4.4}
\end{align*}
$$

The error term $V_{j}$, as already explained in subsection 3.3.2, is given by $U_{j}-\gamma U_{j-1}$ where $U_{j}=\rho U_{j-1}+W_{j}$. When the assumption of $\gamma=\rho$ is invoked, $V_{j}$ becomes equivalent to $W_{j}$ which is white noise by assumption. The equivalence is readily demonstrated by substituting for $U_{j}$ in the
definition of $V_{j}$. The dependent variable $S K$ is the soil test value of $K$ and CK and WK represent potassium applied to corn-1 and wheat, respectively.

Equation (4.2) could have been estimated by applying OLS to a restricted version in which $\theta_{2}=\gamma \lambda_{4}^{-1}$ and $\theta_{3}=\gamma \lambda_{2}^{-1}$. The desired unrestricted coefficients, $\hat{\lambda}_{4}^{-1}$ and $\hat{\lambda}_{2}^{-1}$, could then be obtained as $\hat{\theta}_{2} / \hat{\gamma}$ and $\hat{\theta}_{3} / \hat{\gamma}$, respectively. The results shown in column (c) of Table 4.2 were, however, obtained by applying the nonlinear package in SHAZAM directly to equation (4.2).

The influence of the equality assumption about $\gamma$ and $\rho$ seems to have been observed largely by the constants representing the dummy variables and by $\lambda_{4}$ and $\lambda_{2}$. The values of the former were uniformly increased and so were their t-ratios. The estimated values of $\lambda_{4}$ and $\lambda_{2}$ were also increased. This latter outcome is of particular interest because the two estimates are used to extrapolate the value of the proportionality constant which in turn is used in generating the carryover values of $K$.

The crop coefficients were unaffected both in magnitude and sign. The absolute value of the estimate of the autoregressive parameter was less than one as expected.

In summary, the difference between the two assumptions, $\gamma \neq \rho$ and $\gamma=\rho$, was that in the latter case, the absolute values (but not the relative magnitudes) of the estimated coefficients were increased. The residual variance estimate also increased from 135.05 to 143.36 , a statistically significant increase.

The value of the proportionality constant $\lambda_{k}$ extrapolated from the values of $\lambda_{4}$ and $\lambda_{2}$ given in column (c) of Table (4.2) is 1.6 which is much higher than that derived from column (b). The extrapolation of
$\lambda_{k}$ is done in a manner similar to that of $\lambda_{p}$ as discussed in subsection 4.1.1. The technique for the extrapolation will therefore not be repeated here.

In spite of a better relative fit under the assumption that $\gamma \neq \rho$, the results for the alternative assumption that $\gamma=\rho$ appeared to be more plausible especially in terms of the coefficient t-ratios and the proportion of the applied $K$ carried over from one season to the next. The latter information is obtained as the seasonal estimate of the distributed lag parameter $(\gamma)$. The values of $\gamma$ reported in Table 4.2 are not the seasonal values since the basic time framework in the analysis was a rotation consisting of four growing seasons. The seasonal value is obtained as $\gamma_{k}=4_{\sqrt{|\gamma|}}$ which turns out to be about $77 \%$ for column (c) and only about $38 \%$ for column (b). Both are less than $81 \%$ reported by Lanzer [1978] for the south Brazilian soils.

Subsequent analysis and discussions in this study will proceed on the basis of a potassium carryover function in which it is assumed that $\gamma=\rho$. Hence, attention will now be directed only to column (c) of Table 4.2.

It should be emphasized that the primary reason for estimating the carryover functions is to obtain consistent estimates of the $\lambda$-values. The specification of the function and the technique of estimation ensure that the estimators have the desirable properties. But since different assumptions lead to different specifications it is not uncommon to end up with several sets of estimators. This indeed is true for the case of the potassium carryover function examined here. Three different values of $\lambda_{k}$, two of which are the same magnitude, are given in Table 4.2. The choice among these different estimates must be made on the basis of the
plausibility of the entire set of results. The aim is to go with the assumption which leads to results with the least number of violations to theory and experience already established in crop production. The attempt being made here to rationalize all the estimation results is aimed at establishing the credibility of the totality of the outcome rather than just that of the $\lambda$-values in isolation. The decision to adopt the assumption that $\gamma=\rho$ has been made on the basis of a similar reasoning.

Finally, there is a need for a discussion of the apparent disparity between the influence of crop yields on soil $P$ on the one hand and soil $K$ on the other. The results in Table 4.2 show that the higher the yield of hay/corn-2 and soybean, the lower the following soil test value of K. Wheat and corn-1 had no effect on the soil test at any reasonable significance level.

In the case of $P$, a positive influence of corn-1 on soil $P$ was rationalized as being the result of the gradual decomposition of trash which eventually boosted the soil fertility. The resulting increase in the level of soil $P$ was then registered in the soil test. This explanation cannot be generalized to cover the results for $K$. Potassium is relatively more mobile than $P$. The results for the Indiana soil studied here show that only about $77 \%$ of the applied $K$ is carried over and there is some indication (see the value of $\gamma$ in column (c)) that the figure may ever be much lower. The lower carryover for $K$ implies that either a higher proportion of it is available for the crop or that more of it is lost as compared to P. The former may be more applicable to K. In either case, however, the implication is that the net effect of any production activity is a deficit in the soil level of potassium as indicated
by the negative crop coefficient in column (c).
The results here show that starting with $100 \mathrm{~kg} / \mathrm{ha}$ of available K only about $35 \%$ will be available after four growing seasons (one rotation). With each crop leading to a net loss in the soil fertility by creating a potassium deficit, this nutrient may have to be applied more frequently than $P$ since the decay pattern for the latter shows that only about $30 \%$ of the available form is lost after four growing seasons.

In summary then, the results here show that, for a given block (B) the potassium carryover function is

$$
\begin{align*}
S K_{j}= & B+0.3513\left[\text { SK }_{j-1}+0.1572 C K_{j-1}+0.3135 W_{j}\right] \\
& -0.0030(H A Y / C O R N-2)_{j}=0.0061 \text { SOYBEAN }_{j}+V_{j} \tag{4.5}
\end{align*}
$$

Except for the dummy constant to be substituted by the relevant value from column (c) of Table 4.2, equation (4.5) is similar to the reduced form distributed lag model given in (3.17), Section 3.3. The error term $V_{j}$ is white noise.
4.2 Results for the estimation of crop response functions

### 4.2.1 The polynomial formulations

In this study, the rationale behind the choice of polynomials as alternative formulations in testing the nutrient non-substitution hypothesis has two facets. The first is that of maintaining an apparent tradition of employing polynomials in crop response analysis for whatever reason there may be. The second, and the more important reason, is that the violations of conventional biological theories of plant growth, of which the nutrient non-substitution model is claimed (thus far in this
study) to be free, are most vividly manifested in polynomial specifications.

Among the family of polynomials, the two commonly employed for analyzing crop response to fertilizer are the square root and the quadratic forms. The application of these two functions has been extensive especially over the last three decades. A rigorous exposition of their properties has therefore been deemed unnecessary in this study. ${ }^{\text {// }}$ The essential characteristics and a general critique of polynomials have already been presented in Chapter 2. In this sub-section, the results of estimating the response of one of the crops in the Purdue fertilizer experiment are given. The two polynomials (square root and quadratic forms) analyzed in this study were estimated by means of a multiple regression package provided in the TSP manual.

The crop selected for the response analysis was the first crop of corn designated earlier as corn-1 but which will now be referred to simply as corn since such simplification no longer poses any danger of confusing the crop names. The corn crop was planted in May and harvested in October or November. Broadcast and row fertilizer application rates for this first crop of corn can be gleaned from the Appendix Table A. 2.

To avoid distracting the attention from the major objective of the study, namely to test the validity of the nutrient non-substitution hypothesis, a deliberate effort was made to abstain from cluttering the basic models with a large number of variables. Only the major plant nutrients were considered. However, since the corn crop received nitrogen at

[^6]non-limiting levels, only the influence of the variability in P and K was investigated. No attempt was therefore made to explicitly quantify the contribution of weather and site variables. Such variables can be included in polynomial specifications quite simply but at the expense of increased demands on data. Furthermore, the more variables there are the more obscure is the interpretation of the results of estimation especially for some of the cross-product terms. The question of interpreting the results from approximate functions such as the square root and quadratic polynomials is a controversial one as already mentioned in an earlier chapter. The problem becomes more serious as more variables are crammed in one polynomial function intended for deriving fertilizer recommendations to farmers.

The results discussed below are based on relationships in which the response of corn was considered to be a function of total $P$ and total K. These variables should be viewed as the total supplies of the respective nutrients available to the corn crop during the growing season. In either case, the total nutrient supply was computed as the sum of the applied and residual (carryover) amounts. The latter was obtained by multiplying the soil test index by the $\lambda$-value derived from Section 4.1. The data used in estimating the response of corn to total P and K are presented in Appendix Table A.3. The yield and applied fertilizer observations have been converted to quintals/ha and the soil tests are in $\mu \mathrm{g} / \mathrm{g}$. Due to the presence of gaps in the soil test data, only the observations for six years were used in analyzing the response of the corn crop. The corn data span the years 1960 through 1966. The corresPonding soil test values are for the years 1959 through 1965. The lag is accounted for by the fact that soil sampling was taken prior to the
growing season, hence the soil test value used for constructing the nutrient carryover for the 1960 corn crop is that of 1959.
$A$ value of $\lambda_{p}=0.1226$ was used to construct total $P$. This figure was obtained by dividing that given in Table 4.1 column (c) by 100; $\lambda_{p}$ as given in that Table is in $\mathrm{kg} / \mathrm{ha}$ of P whereas the data used in estimating the response of corn is in quintals/hectare. The value of $\lambda_{k}$ was taken as 0.016 . This is $1 / 100$ times the figure given in column (c) of Table 4.2

To facilitate direct comparisons of results from the estimation of the square root and quadratic functions with those of the non-substitution model, the corn yield data as given in Appendix Table A. 3 were further divided by 100. The reason for doing this is that the yield data measured in quintals/hectare did not meet the scaling requirements in the mathematical programming technique used for estimating the non-substitution model.

As in the case of the carryover functions discussed in the previous section, the effects of the different blocks were accounted for by four dummy variables $S_{i}(i=1,4)$. The square root function estimated was of the form

$$
\begin{align*}
\text { CORN }= & S_{1}+S_{2}+S_{3}+S_{4}+\gamma_{1} P T+\gamma_{2}(P T)^{\frac{1}{2}}+\gamma_{3}(P T K T)^{\frac{1}{2}} \\
& +\gamma_{4}(K T)^{\frac{1}{2}}+\gamma_{5} K T+\varepsilon_{1} \tag{4.3}
\end{align*}
$$

where PT is total supply of P given by $0.1226 \mathrm{SP}+\mathrm{CP}$ and KT is the total supply of $K$ given by 0.016 SK + CK. SP and SK are the soil test values for $P$ and $K$ whereas $C P$ and $C K$ are the applied amounts for the respective nutrients.

The observations used in estimating equation (4.3) were the 44 treatment values for one growing season (22 for each replicate). The
individual error terms $\varepsilon_{1 t}$ ( $t$ denoting treatment) were assumed to be iid with zero mean and a common variance $\sigma_{1}^{2}$. As already explained in a previous section and as can be deduced from Table 3.1, the same iid assumption can be extended to the functions pooling more than one growing season; the observations were blockwise independent.

In equation (4.3), the coefficients $\gamma_{1}$ and $\gamma_{5}$ measure the linear effects of PT and KT on the yield of corn whereas $\gamma_{2}$ and $\gamma_{4}$ are measures of the degree of curvature of the function. The coefficient $\gamma_{3}$ on the other hand measures the effects of interaction between PT and KT.

Table 4.3 summarizes the results for the estimation of the single year square root functions for the years 1960 through 1966 . The table also includes the results for the estimation of a function pooling the seven years. The table is self-explanatory but a few comments are in order. First, it should be noted that no interaction effects between P and K were registered during the years of investigation. Secondly, the coefficients of the linear and the square root terms had the expected signs; negative for the former and positive for the latter. The third comment is that a casual survey of the table reveals that there was variation in response of corn from year to year. Of particular interest is the fact that although the linear and the square root terms for PT . had significant coefficients in the pooled function, phosphorus had a significant contribution in only three out of the seven years reported in Table 4.3. The t-ratios for the coefficients of the linear and the square root terms for potassium on the other hand suggest that this nutrient had a significant contribution to the response of corn in each of the seven years investigated. This differential contribution of $P$ and $K$ over the years could be due to the differences in the mobility

Table 4.3 Results for the Estimation of the Square Root Function ( $10^{2} \mathrm{~kg} / \mathrm{ha}$ )

${ }^{\mathrm{a}}$ The figures in parenthesis are the coefficient t-ratios. ${ }^{\mathrm{b}}$ Short-hand for (PT) $)^{\frac{1}{2}}$, (PTKT) ${ }^{\frac{1}{2}}$ and (KT) $)^{\frac{1}{2}}$, respectively. $\mathrm{C}_{\text {Residual }}$ variance estimate (RVE) $=$ SSE/degrees of freedom.

亏．
and hence in the residual values of these nutrients. It is likely that by 1960, the soil level of the two nutrients was such that only K was limiting growth. The changes in weather and soil conditions and in amounts of $P$ permanently taken out of the field may at times have reduced the availability of this nutrient to levels which could limit growth as indicated in the years 1963, 1965, and 1966.

The variation in crop response from year to year is an important subject area particularly when it comes to the choice of the combination of years of data to be used in estimating the response function which is consequently employed in deriving optimal fertilization strategies. The results in Table 4.3 show that the pooled function gives the false impression that $P^{\prime}$ had a significant contribution to the response of corn. It is therefore likely that using the pooled function may lead to recommendations requiring annual addition of P whereas the single year functions suggest a lower frequency of application.

A look at the row labeled RVE (residual variance estimate) in Table 4.3 also indicates that the assumption of constant variance from year to year may not be valid for the pooled model. This would imply that even though the coefficient estimates from the pooled model are unbiased they are not minimum variance estimators. To get more efficient (maximum likelihood) estimators under the assumption of blockwise shifting variance, one would have to resort to generalized least squares (GLS) techniques. In order to stay within the boundaries of the stated objectives of this study, the application of GLS to the pooled models and the topics relating to yield stochasticity were not considered. However, it is apparent from Table 4.3 that these are important areas where further analysis of the present data set could yield significant payoff.

The final comment to be made about Table 4.3 concerns the reported $R^{2}$ values which averaged about 0.55 for the single year models. For time series observations this figure borders on the low side. In the present case, however, it is possible to rationalize the low $\mathrm{R}^{2}$ values on grounds of the cross-sectional nature of the data used and the fact that important sources of variation in the corn yield may have been omitted. Heading the list of such excluded factors are the weather and site variables. It is also possible that the two-stage approach employed in deriving total nutrient supplies (PT and KT) based on extrapolated $\lambda$-values may have led to some loss in the precision of coefficient estimates. Since estimates of PT and KT could not be obtained directly, such a loss in efficiency was unavoidable. It must be emphasized here, however, that the relative performance of the individual models is of no consequence in the anticipated tests of hypotheses regarding the appropriateness of functional specification. Except for the manipulations required by the different specifications, the same independent variables PT and KT were maintained in all the models estimated. It follows that statistical differences between results of a given pair of models must be due to differences in specification.

The results for the estimation of the quadratic polynomial summarized in Table 4.4 indicate a seasonal response pattern similar to that obtained from the square root formulation. The quadratic equation estimated can be obtained from (4.3) by replacing the square root terms with the quadratic terms (PT) ${ }^{2}$, PTKT and $(K T)^{2}$. The symbols used for the block durmy variables are $Q_{i}(i=1,4)$ in the case of the quadratic function. As in the square root form, the estimated coefficients for the block dummy variables were highly significant at any reasonable level of

Table 4.4 Results for the Estimation of the

|  | 1960 | 1961 | 1962 | 1963 |
| :---: | :---: | :---: | :---: | :---: |
| $\mathrm{Q}_{1}$ | $\begin{gathered} 74.4564 \\ (17.3)^{a} \end{gathered}$ | -- | -- | -- |
| $\mathrm{Q}_{2}$ | -- | $\begin{aligned} & 62.9909 \\ & (12.6) \end{aligned}$ | -- | -- |
| $Q_{3}$ | -- | -- | $\begin{aligned} & 72.7921 \\ & (9.1) \end{aligned}$ | -- |
| $Q_{4}$ | -- | -- | -- | $\begin{aligned} & 80.1309 \\ & (20.8) \end{aligned}$ |
| PT | $\begin{aligned} & 3.4964 \\ & (2.3) \end{aligned}$ | $\begin{aligned} & 1.0906 \\ & (0.5) \end{aligned}$ | $\begin{aligned} & 0.8486 \\ & (0.3) \end{aligned}$ | $\begin{aligned} & 4.9988 \\ & (3.5) \end{aligned}$ |
| $(P T)^{2}$ | $\begin{aligned} & -0.2751 \\ & (1.6) \end{aligned}$ | $\begin{gathered} -0.1190 \\ (0.5) \end{gathered}$ | $\begin{gathered} -0.2976 \\ (0.9) \end{gathered}$ | $\begin{aligned} & -0.5897 \\ & (3.4) \end{aligned}$ |
| PTKT | $\begin{aligned} & -0.0948 \\ & (1.0) \end{aligned}$ | $\begin{gathered} -0.0614 \\ (0.4) \end{gathered}$ | $\begin{aligned} & 0.2531 \\ & (1.3) \end{aligned}$ | $\begin{aligned} & 0.0457 \\ & (0.4) \end{aligned}$ |
| $(K T)^{2}$ | $\begin{aligned} & -0.5514 \\ & (3.2) \end{aligned}$ | $\begin{gathered} -1.0771 \\ (4.5) \end{gathered}$ | $\begin{gathered} -1.2344 \\ (3.9) \end{gathered}$ | $\begin{aligned} & -0.5803 \\ & (3.3) \end{aligned}$ |
| KT | $\begin{aligned} & 6.8245 \\ & (3.6) \end{aligned}$ | $\begin{gathered} 10.8524 \\ (4.7) \end{gathered}$ | $\begin{gathered} 11.8918 \\ (3.5) \end{gathered}$ | $\begin{aligned} & 6.1036 \\ & (3.4) \end{aligned}$ |
| $R^{2}$ | 0.5183 | 0.5537 | 0.4944 | 0.5758 |
| RVE | 19.9980 | 37.3045 | 68.7537 | 20.7131 |

Quadratic Function ( $100 \mathrm{~kg} / \mathrm{ha}$ )

| 1964 | 1965 | 1966 | Pooled |
| :---: | :---: | :---: | :---: |
| 73.4647 <br> $(14.8)$ | -- | -- | 72.1627 |
|  |  |  | $(26.4)$ |
| -- | 63.4734 | -- | 65.5730 |
|  | $(13.8)$ |  | $(23.9)$ |
| -- | -- | 53.3757 | 62.7537 |
|  |  | $(8.8)$ | $(22.9)$ |
| -- | - | -- | 77.8333 |
|  |  |  | $(27.0)$ |
| 2.6600 | 4.6778 | 4.6047 | 3.9255 |
| $(1.4)$ | $(2.7)$ | $(2.4)$ | $(4.1)$ |
| -0.2427 | -0.4527 | -0.6137 | -0.4619 |
| $(1.1)$ | $(2.4)$ | $(3.0)$ | $(4.2)$ |
| -0.0089 | 0.0271 | 0.0612 | 0.0524 |
| $(0.1)$ | $(0.2)$ | $(0.4)$ | $(0.8)$ |
| -0.7140 | -0.8012 | -0.7075 | -0.7660 |
| $(3.1)$ | $(3.8)$ | $(2.8)$ | $(6.7)$ |
| 6.9921 | 8.4741 | 7.7125 | 8.0087 |
| $(3.0)$ | $(4.0)$ | $(2.9)$ | $(6.8)$ |
| 0.3177 | 0.5705 | 0.5229 | 0.4583 |
| 35.8294 | 27.2166 | 42.5729 | 546.9579 |


probability for a type I error. Not only were the coefficients significant, they were also relatively large in absolute terms, a fact which is probably due to the omission of some variables as already noted.

In terms of data "fit," as judged from the $R^{2}$ and the residual variance estimate, the square root specification appears to have a slight edge over the quadratic form. No definitive statements can be made on the subject of discrimination between these two specifications without carrying out proper statistical tests. Such tests will be done later on but with nutrient non-substitution as the alternative hypothesis. Due to the popularity of the two estimated polynomial functions and because the results show no definite superiority for either one in the absence of more rigorous tests, both specifications will be tested against the non-substitution model.

### 4.2.2 The nutrient non-substitution model

The nutrient non-substitution model was estimated in the form given in equation (3.32) of Section 3.4. Before doing so, however, estimates of the yield plateaus for the individual years had to be provided. In addition, the knots needed in estimating the spline functions were unknown and hence had to be estimated.

The realized yield plateau or maximum (A) for a given crop variety and location depends, among other things, on the prevailing weather conditions. Under idealized conditions, the biological maximum yield (M) for a given crop variety will be a constant depending on the variety's genetic potential. In this study, the realized maximum was estimated as the average of the top five corn yields. For example, the maximum
yield for the 1960 corn data was taken as 103.0 quintals/ha. This, admittedly, is an arbitrary way of computing the yield maximum but the figure is not only a reflection of the prevailing weather conditions but also of the biological maximum, M. It is likely that there is a direct relationship between $M$ and the estimate $A$ of the realized maximum used here. Lanzer [1978] postulated that the yield maximum given by $A=\operatorname{Max}\left[y_{1} \leq y_{2} \leq \cdots . \leq y_{N}\right]$, an order statistic, had a proportional relationship with the constant $M$. He estimated the constant of proportionality simultaneously with the other parameters of the model using a non-linear algorithm. The framework of estimation adopted here precludes such an approach. For example, for a single year model under the assumption that $M=\alpha \hat{A}$ an exercise of parameterizing over different values of $\alpha$ in the range [0.5-1.6] merely shifts more or less weight to the other coefficients of the model leaving the error sum of squares unaffected. For more than one year, different combinations of $\hat{A}$ and $\alpha$ values may lead to different fits for the data but such fine-tuning was deemed to be of marginal value. Hence, A was taken as the average of the top five observations for a given year of yield data for corn.

The non-substitution model estimated had a nonlinear objective function and a set of linear and nonlinear constraints. As already stated, the computer program used was MINOS/AUGMENTED which will be referred to simply as MINOS. The knots were treated essentially as unknown and had to be searched for by repeated estimations of the model. Before proceeding to explain the search procedure for the knots, a description will be given of how the problem was set up for estimation by the MINOS program.
following problem-specific information: CALCFG, CALCON, SPECS, and the MPS. These terms are explained in the MINOS User's Guide and MINOS/AUGMENTED User's manual co-authored by Murtagh and Saunders [1977 and 1980]. The explanations given below are only for continuity of exposition.

CALCFG is the fortran subroutine used to compute the objective function $f(X)$ and the gradient vector $g(x) . X$ is the set of all nonlinear $\left(X_{N L}\right)$ and linear ( $X_{L}$ ) variables. For the problem at hand, as can be seen from equation (3.32), $X_{N L}=(S K S P \varepsilon)$ and $X_{L}=\left(\mu Z K_{j} Z P_{j}\right)$ in which $S_{1}$ and $S_{2}$ have been substituted by $S K$ and $S P$ and $\beta_{1}$ and $\beta_{2}$ by $Z K_{j}$ and $Z P_{j} . K$ and $P$ are the two nutrients and $j=1,2,3, \ldots$ is the number of columns of the matrices $Z K$ and $Z P$ which have now replaced $Z_{1}$ and $Z_{2}$ of equation (3.32). The nonlinear variables SK and SP appear only in the nonlinear constraints and not in the objective function. Therefore the function $f(X)$ is $\frac{1}{2} \varepsilon^{\prime} \varepsilon$ hence $g(X)=\varepsilon$ and these go into CALCFG given in the Appendix A. 4.

CALCON is a subroutine similar to CALCFG but designed to compute the nonlinear constraints and their gradients. For this problem, the nonlinear constraint variables are SK and SP. Each treatment level has the accompanying quadratic constraint of the form $S K_{t} S P_{t}=0, S K_{t} \geqq 0$ and $S P_{t} \geqq 0$. The constraint gradients are therefore given as $g\left(S K_{t}\right)=S P_{t}$ and $g\left(S P_{t}\right)=S K_{t}$ for all $t$. A fortran version of CALCON is given in Appendix A. 5.

SPECS stands for specifications. This is the section for stating the dimensions of the problem. All the non-substitution problems estimated by MINOS program had 4 N rows and ( $4 \mathrm{~N}+2 \mathrm{~J}$ ) columns where $\mathrm{N}=44$ is the number of observations in one year. $J$ is the number of columns of the matrix $Z K$ which is also the dimension of the matrix ZP (remember
that $Z K$ and $Z P$ replaced $Z_{1}$ and $Z_{2}$, respectively, in equation (3.32)). In each problem, there were $N$ nonlinear constraints, $2 N$ nonlinear (jacobian) variables corresponding to SK and SP and $3 N$ nonlinear objective variables corresponding to $X_{N L}=(S K S P \varepsilon)$. Notice that in CALCFG, only the gradients for $\varepsilon_{t}$ are computed since $S K$ and $S P$ have no value in the objective function.

The SPECS file is listed every time in the MINOS printout. A printout for the 1960 problem is reproduced in Appendix Table A.7. Notice that the rows and columns specified in the SPECS are overestimates of the actual figures. This is a program requirement.

General bounds on all variables can also be included in the SPECS. For example, it was found that the problem setup was such that after appropriate scaling of the data, all coefficient estimates would lie below the largest value of the right hand side (RHS) vector - corn yield data for the printout given in Appendix Table A.7. It was therefore decided that a value slightly larger than the largest of the RHS values would be placed in the SPECS as the upper bound. A properly selected upper bound restricts the parameter space and hence reduces the number of major iterations.

The SPECS may also include tolerances (used in the algorithm) which the user desires to set at levels other than their default values. The latter are listed together with all the other parameters in the section of the printout (Table A.7) bearing the title PARAMETERS. The estimations carried out here left most of the parameters at their default values.

The mathematical programming section (MPS) is used to specify the locations of the constraint gradients (by dummy variables) and the elements of the matrix of coefficients for the linear constraints. In
the present case, the relevant matrices are $A_{1}$ and $A_{2}$ defined in equation (3.32). The MPS file for the 1960 problem is reproduced in Appendix Table A. 6.

The MPS file corresponds in a rather straightforward manner to the constraint layout in equation (3.32). There are four sets of rows (all of the equality type) with the names ROWS, ROWY, ROWZK and ROWZP. Each set has $N$ row elements giving a total of $4 N$ rows. There are six sets of columns named COLSK, COLSP, COLEE, COLMU, CLZK ${ }_{j}$ and CLZP ${ }_{j}$ corresponding to the nonlinear and linear variables of the model. The last two sets contain the coefficients for the spline functions for the two nutrients $K$ and $P$. The actual number of coefficients depends on the number of knots, a subject which will be discussed shortly. For the 1960 problem, the coefficients corresponding to the slope of the first segment of the spline curve ( $\beta_{i 0}$ in Fig. 3.1) are CLZK2 and CLZP2. The coefficients representing the change in slope are named CLZK3 and CLZP3. The additional coefficients CLZK1 and CLZP1 represent the block dummies and for the single years appear as intercepts. They should not, however, be interpreted as intercepts since the models employ total supplies of the nutrients. Ideally, the spline functions should be forced through the origin in order to conform to the concept of nutrient essentiality.

The only nonzero entries in the RHS section appear in ROWY. These are the corn yield data after the transformations discussed in Appendix A. 1.

The last part of the MPS file is the BOUNDS section. Notice that CLZK3 and CLZP3 are constrained to be nonpositive thus ensuring concavity of the spline functions. The error term $\varepsilon$ is unrestricted hence the corresponding column COLEE is coded in the BOUNDS section as free
(FR). The other variables SK, SP, MU, ZK1, ZK2, ZP1, and ZP2 are not entered in the BOUNDS section explicitly and are therefore coded, by default, as G i.e. $\geqq$.

An additional restriction was added to the basic framework summarized thus far in the MPS. This was the requirement of a plateau surface which from Fig. 3.1 implies that

$$
\begin{equation*}
\beta_{i}=\sum_{\tau=0}^{j} \beta_{i \tau}=0 \tag{4.4}
\end{equation*}
$$

where $i$ represents the two nutrients and $j+1$ is the total number of knots. The summation in (4.4) gives the slope of the spline function at the $j^{\text {th }}$ knot. In keeping with the variable names introduced for the MPS file presented in Appendix Table A.6, the plateau restriction requires that

$$
\text { and } \begin{align*}
Z K 2+Z K 3 & =0  \tag{4.5}\\
Z P 2+Z P 3 & =0 \tag{4.6}
\end{align*}
$$

noting that both ZK3 and ZP3 are nonpositive by restriction. The additional restrictions (4.5) and (4.6) were accommodated in the MPS file by introducing two extra rows: ROWKS and ROWPS.

For the spline function approximations to the single nutrient response curves, only one effective knot was used, that is, not counting the final knot which for both PT and KT was taken as $830 \mathrm{~kg} / \mathrm{ha}$. This amounted to estimating simple models of the linear and plateau type therefore the spline function approximations had only two segments (refer back to Fig. 3.1).

The last knot, for either nutrient, was not difficult to locate since it had to be chosen so as to exceed the largest recorded observation. The single effective knot referred to above was not as easy to pick.

Generally, a knot is supposed to mark the location of a change in the slope of the function. Such locations were not easily discernible by merely glancing at the observations or the scatter diagram of yield and either PT or KT. The knots were therefore taken as unknown variables which had to be chosen so as to minimize the value of the objective function as given in equation (3.32). The simplistic nature of the linear and plateau type approximations adopted here made this task rather easy. First, the scatter diagrams were used to select initial ranges of the knots for PT and KT. For example, the range for the 1960 data for both nutrients was taken as $100-350 \mathrm{~kg} / \mathrm{ha}$. Values of the objective function corresponding to these extreme points were computed and then different pivots ( $k$ ) within that range were tried. The points which minimized the error sum of squares, $S(k)$, were taken as the final knots. The final knots for the 1960 data were $155 \mathrm{~kg} / \mathrm{ha}$ of PT and $155 \mathrm{~kg} / \mathrm{ha}$ of KT . The knots corresponding to the minimum points of $S(k)$ for the years studied are enclosed in square brackets on the k-axis of Fig. 4.2. As can be seen from the figure, only a few points were tried for each year before locating the minimum point of $S(k)$.

If the range between zero and the final knot is large, it is possible to include an additional knot in order to improve the "fit." The knots . indicated in Fig. 4.1 were relatively low compared to the final knot which was fixed at $830 \mathrm{~kg} / \mathrm{ha}$. Hence, there was little to be gained by having additional knots. The spline functions estimated, therefore, had only two segments: a linear portion and a plateau.

The computer printouts listing the iteration $\log$ and the entire solution for any given year is rather lengthy. A typical printout for the 1960 problem is presented in Appendix Table A.7. The column entitled

Figure 4.2 The relationship between the (non-substitution model) residual sum of squares $S(k)$ and the knots ( $k$ )


ACTIVITY in Section 2 of the solution gives the coefficient estimates for the variables $S K, S P, \varepsilon, \mu, Z K_{j}$ and $\mathrm{ZP}_{j}(j=1,3)$, in that order. The objective value is given in the section preceding the solution listing and the required error sum of squares, $S(\varepsilon)$, is twice that value since the minimized function $f(X)=\frac{1}{2} \varepsilon{ }^{\prime} \varepsilon$.

This [1960] problem was given a "warm start" by providing basis data to define an initial solution. Convergence was therefore achieved after only 4 major iterations. The initial basis data were saved from trial runs operating at different knots. It was essential to start a new problem from an old basis, one saved from another problem, even if the solution to the source problem fell short of expectation. The desired improvements were then achieved in a stepwise manner.

The estimated spline function parameters for the 1960 problem are summarized in Table 4.5 and the estimated nutrient non-substitution model for corn response is therefore given by

$$
\begin{align*}
\operatorname{CORN}= & 103.0 \operatorname{Min}\{0.6829+0.1743 \mathrm{KT}, 0.8197+0.0909 \mathrm{PT}\}: \\
& \mathrm{KT}, \mathrm{PT} \leqq 1.55 \text { quintals } / \mathrm{ha} \\
= & 0: \mathrm{KT}, \mathrm{PT} \in[1.55-8.30] \text { quintals } / \mathrm{ha} \tag{4.7}
\end{align*}
$$

where KT and PT represent total nutrient supply for $K$ and $P$, respectively, and the estimate of the maximum yield ( $\hat{A}$ ) is given as 103 quintals/ha. The last knot fixed at 8.30 quintals/ha was considerably larger than the estimated knot, 1.55 quintals/ha. The range [1.55-8.30] quintals/ ha therefore represents a region over which the response of the corn crop was not limited by either of the nutrients studied here.

The range [0-1.55] quintals/ha covers the region over which there is a positive response to both nutrients as given by the coefficients of KT and PT in equation (4.7). The asymptotic standard errors for

Yield Maximum
Potassium
Block dummy variable
Relative yield function slope for $\mathrm{KT}=[0-1.55]$ quintals $/ \mathrm{ha}$
Change in slope at $\mathrm{KT}=1.55$ quintals/ha

0.6829
hosphorus
Block dummy variable

## ZP1

0.8197

Relative yield function slope for $\mathrm{PT}=[0-1.55]$ quintals/ha
Change in slope at $\mathrm{PT}=1.55$ quintals $/$ ha RESIDUAL VARIANCE ${ }^{\text {b }}$
${ }^{\text {a }}$ Measurements are in $100 \mathrm{~kg} / \mathrm{ha}$. Since the RHS values (yield data) were further divided by 100 , the reported coefficient estimates are those in Appendix Table A. 7 multiplied by 100 then divided by the estimated yield maximum (A). The rationale behind this should be clear from equation (3.31).
${ }^{\mathrm{b}}$ The residual variance estimate, $R V E=S S R /(N-K)$. In the non-substitution single year models, $N=44$ and $K=6$. The residual sum of squares (SSR) is obtained by multiplying the estimated objective value in Appendix Table A. 7 by 2 and further multiplying by $100^{2}$.
these coefficients were not computed but the results summarized in equation (4.7) and Table 4.5 seem to confirm those of the square root and quadratic formulations, namely that the corn crop was over supplied with K and P in 1960. For the polynomial specifications, since the coefficient standard errors were provided, it was readily determined that, on the average, there was a significant response to $K$ but not to P. The non-substitution model estimated here portrays the same information but in a different way. It says that positive response to both nutrients was restricted only to the region below 1.55 quintals/ha with the one exception that larger yield increments would be obtained by additional units of K rather than of P . This kind of information, together with soil test results can form the basis for quick fertilization guidelines to farmers.

In spite of its simplicity, equation (4.7) is powerful enough in the sense that it contains the basic information that may be required by most of the farmers. The required information is about the returns to the scarce resources (especially capital) diverted to crop production. In particular, there is a need for information relating to rates of fertilizer application. The estimated model can be used as already noted earlier (see equation (3.27)) to derive long term optimal fertilization strategies. Such optimization techniques tailored to specific soils and weather conditions may be beyond the budget lines of most of the agricultural extension systems especially in the less developed economies. All that an extension agent may need, therefore, are the soil test results and a simple formula such as (4.7) relevant for a given soil type and remembering that $K T=K+0.0159 \mathrm{SK}$ and $\mathrm{PT}=\mathrm{P}+0.1226 \mathrm{SP}$ where $K$ and $P$ are the required application rates and $S K$ and $S P$ are in
soil test units of potassium and phosphorus, respectively. The knot given as 1.55 quintals/ha of either $K$ or $P$ marks the critical level of total supply of these nutrients. Supplies beyond the critical level are considered as overfertilization of the crop and hence a waste.

The basic differences between the polynomials and the simple model in (4.7) from the point of view of application, are the more involved algebraic manipulations encountered in the former. There is also the risk that the polynomials will overestimate the critical level of nutrient supply corresponding to the maximum possible yield. The estimated square root function for 1960, for example, gives the levels of $K$ and $P$ corresponding to the maximum yield as 6.319 quintals/ha and 5.509 quintals/ha respectively. The corresponding figures for the quadratic function are 5.727 and 5.368 . These two sets of figures are, for all practical purposes, similar. The estimated maximum yields for the two models are, however, not of the same magnitude. The quadratic tops at a yield level of 121.487 quintals/ha whereas the square root form is flatter, reaching its peak at 101.083 quintals/ha. The largest recorded corn yield for the 1960 data was 105.34 quintals/ha. It is therefore easy to see that, in comparison to the estimated non-substitution model and the information on actual crop yield, the quadratic had a tendency to overstate not only the region of positive response to total nutrient supply, but also the corresponding yield. The yield maximum estimated by the quadratic was more than 4 standard deviations in excess of the yield maximum ( $\hat{A}$ ) used in the non-substitution model and more than 3 standard deviations above the maximum recorded yield of corn in 1960.

The square root formulation, in comparison to the non-substitution
model, exaggerated the region of positive crop response but otherwise had a yield maximum close to the approximation $\hat{A}$ used in the latter specification.

As can be deduced from Table 4.5, the standard deviation for corn yield estimated by the nutrient non-substitution model was $428.9 \mathrm{~kg} / \mathrm{ha}$ which, using a conversion factor of 25.4 kg for every bushel of shelled corn, is equivalent to 16.9 bu/ha. The corresponding figures for the square root and quadratic polynomials were 17.2 bu/ha and 17.6 bu/ha, respectively. These figures show that, purely on the basis of the standard error of the regression, the simple linear and plateau model performed no worse than the popular polynomials. The same conclusion has been stated by other researchers with different methodologies of estimating models based on Liebig's law of the minimum (see Chapter 2). Such comparisons, however, may not be statistically valid as a basis for making unambiguous judgments with regard to the choice of functional specification. Indeed, the present case bears that fact out since there seems to be a deadlock between the non-substitution model and the polynomial specifications in terms of the standard error of the regression after correcting for degrees of freedom (the square root of the reported RVE).

In spite of the apparent similarity as judged from data fitting, the three models imply different fertilization strategies in view of the differences already mentioned above. Economic criteria as a basis of discriminating among these models should therefore be precluded not only because of the pervasive manner in which the biases inherent in the models under comparison are manifested in such criteria but also because such a procedure may not be statistically appropriate. It
would not be statistically appropriate because, as already discussed in Section 2.2, the models are not different transformations of a more general specification and hence do not lend themselves easily to direct comparison. Furthermore, as with the case of using RVE and $R^{2}$, there is the presumption that the real specification is known. In crop response analysis, comparisons which presuppose knowledge of the actual functional relationship may not be justified.

In this study, each model will be viewed as constituting a hypothesis, the validity of which has to be tested against specified alternatives. A priori, there will be no maintained hypothesis. However, there will be special interest on the performance of the nutrient non-substitution hypothesis in view of its conformity to conventional theories of plant growth. Hence, this hypothesis will be tested against the polynomial approximations. The latter will in turn be tested against the nutrient non-substitution hypothesis. Tests will be done not only for the models estimated from the 1960 data but also for those derived from the years 1963, and 1965. Non-nested hypothesis testing procedures will be employed. Three different test statistics will be used: the Cox-Pesaran (CP) statistic, the linearized CP statistic and the C-statistic. Descriptions of these test statistics have already been given in Section 3.5.

The estimated spline functions for the other selected years are presented in Tables 4.6 and 4.7. For 1963, the final knot was found to be about half the estimated values for 1960 and 1965 as can be seen from Figure 4.2. In terms of RVE, the 1963 and 1965 results indicated that the non-substitution model fitted the data better than either the square root or the quadratic specifications. This is in contrast to

Potassium
Block dummy variable
Relative yield function slope for $K T=[0-0.85]$ quintals $/$ ha
Change in slope at $K T=0.85$ quintals $/$ ha
Phosphorus
Block dummy variable
ZP1
0.6383

Relative yield function slope for PT $=[0-0.85]$ quintals/ha
Change in slope at $\mathrm{PT}=0.85$ quintals/ha
RESIDUAL VARIANCE

| $Z P 2\left(\beta_{P 0}\right)$ | 0.4711 |
| :--- | ---: |
| $Z P 3\left(\beta_{P 1}\right)$ | -0.4711 |
| RVE | 14.1018 |

Table 4.7 Estimated Spline Functions - 1965

## Yield Maximum

Potassium
Block dummy variable
Relative yield function slope for $\mathrm{KT}=[0-1.65]$ quintals $/ \mathrm{ha}$
Block dummy variable
Relative yield function slope for $\mathrm{KT}=[0-1.65]$ quintals $/ \mathrm{ha}$
Change in slope at $K T=1.65$ quintals $/$ ha
Phosphorus
Block dummy variable
Relative yield function slope for PT $=[0-1.65]$ quintals/ha
Change in slope at PT $=1.65$ quintals/ha
RESIDUAL VARIANCE

A
93.6500

| ZK1 | 0.5071 |
| :--- | ---: |
| ZK2 ( $\left.{ }_{K O}\right)$ | 0.4667 |
| ZK3 ( $\left.{ }_{K K 1}\right)$ | -0.4667 |

ZP1
0.7502
$Z P 2\left(B_{P O}\right)$
0.1521

ZP3 ( $\beta_{P 1}$ )
$-0.1521$
RVE
18.4974
the 1960 results which indicated only a marginal advantage of the nonsubstitution model over the other specifications. It is hoped that the intended non-nested hypothesis tests will bring out a clearer distinction between the two sets of models being tested than is possible when comparisons are based only on RVE.

No more comparisons of the models on the basis of their relative performance criteria will be undertaken here. Furthermore, since the tables (4.6 and 4.7) are self-explanatory (the derivation of coefficients being similar to that of Table 4.5), the 1963 and 1965 results will not be elaborated. The complete results of estimation for the non-substitution models for these two years are presented in Appendix Tables A. 8 and A.9.

### 4.3 Results and Discussion of Hypothesis Tests

The results for the hypothesis tests are discussed in this section. Three specifications were tested: the non-substitution model and the square root and the quadratic polynomials. The polynomials were not tested against each other since the basic interest was to test them against the non-substitution formulation. Therefore, the results presented in this section are for the following hypothesis tests, abbreviated as indicated:

| $H_{0}$ Tested Against $\mathrm{H}_{1}$ | Abbreviation |
| :--- | :---: |
| Non-substitution (NS) against Square root (SR) | NS/SR |
| Square root against Non-substitution | SR/NS |
| Non-substitution against Quadratic (Q) | NS/Q |
| Quadratic against Non-substitution | Q/NS |

For the three years studied, this gives 12 tests; 6 against the polynomial specifications (NS maintained) and 6 against the non-substitution specification (either SR or Q maintained). Only the abbreviations are used in subsequent tables.

Basically, there were two testing procedures: the C-test and the Cox-Pesaran (CP) approach. A third test was obtained by linearizing the CP statistic, a relatively easy operation. The C-test was used as an easy and preliminary test. Its results will be compared with those of the CP procedure. The outcome of such comparisons may give some indications as to the consistency of the C-statistic as a quick means of testing model specification.

As stated in Section 3.5, the linearized CP statistic is more conservative at rejecting the null hypothesis than the CP statistic when the alternative is fitting much better than expected. It is therefore used as an adjustment to the CP statistic to counteract the unexpected superiority of the alternative hypothesis.

The results discussed first are for the C-test. For a pair of hypotheses, the desired statistics were obtained by estimating the two equations:

$$
\begin{array}{ll} 
& y-\hat{\mu}=\theta_{0}\left(x^{\top} \hat{\gamma}-\hat{\mu}\right)+\varepsilon_{0} \\
\text { and } y-x^{\top} \hat{\gamma}=\theta_{1}\left(\hat{\mu}-x^{\top} \hat{\gamma}\right)+\varepsilon_{1} \tag{4.8.2}
\end{array}
$$

The first equation is a reproduction of equation (3.35) and is used to
test the non-substitution model against a polynomial formulation. All the symbols in (4.8.1) are vector valued and retain the same interpretation given in Section 3.5. It is, however, worth noting that while $\hat{\mu}$ is the expected yield estimated by the non-substitution model, $x^{\top} \hat{\gamma}$ symbolizes the expected yield as estimated by either the square root or the quadratic, depending on which is in the alternative hypothesis. When the non-substitution hypothesis is maintained, as in (4.8.1), the interest is on an estimate of $\theta_{0}$ and its standard error. These two estimates give the t-ratio which is the C-statistic used to test the validity of the maintained hypothesis.

Whereas the first equation tests the validity of the non-substitution hypothesis, the second tests the polynomial approximations (SR or Q) with the nutrient non-substitution hypothesis as an alternative.

The expected yield estimated by the non-substitution model was given by $\hat{\mu}=2 \hat{K}_{K}$ or $\hat{\mu}=Z P_{\beta_{p}}$ depending on whether potassium (K) or phosphorus $(P)$ was the growth limiting nutrient and where the right hand expressions are the estimated spline functions. In (4.8.1), $\theta_{0}$ is a nesting parameter, this role being obvious when the equation is written out as $y=\left(1-\theta_{0}\right) \hat{\mu}+\theta_{0} x^{\top} \hat{\gamma}$. Substituting $\hat{\beta}^{\prime}=\left(\hat{\beta}_{K} \hat{\beta}_{p}\right)$ for $\beta^{\prime}=\left(\beta_{K} \beta_{p}\right)$ and $\hat{\gamma}$ for $\gamma$ is just one way of parametric identification of the resulting comprehensive model. Any other estimators could have been used provided they were consistent for $\beta$ and $\gamma$. The resulting estimate of $\theta_{0}$ is therefore conditional on $\beta$ and hence not unique. Its $t$-ratio is, however, unique and can thus be used to test the validity of the tested hypotheses.

Asymptotically, the estimate of $\theta_{0}$ in (4.8.1) will converge to unity when the polynomial specification is valid. Alternatively, it will converge to zero when the tested hypothesis is true. These facts may be
used to conclude that if $\theta_{0}$ is significantly different from one then the alternative hypothesis, in this case the polynomial, is not supported by the data. However, in view of the foregoing discussion about the estimated (absolute) value of $\hat{\theta}_{0}$ it was necessary to reverse the roles of the hypotheses under test and hence the need for the second equation in (4.8). The tests were then based on the estimated t-statistics for $\hat{\theta}_{0}$ and $\hat{\theta}_{1}$.

The C-tests are easy to perform and the resulting statistics will be used in this study both as preliminary crude tests and also for the purposes of comparison with the Cox-Pesaran (CP) procedure. The latter yields a statistic which is asymptotically equal to the negative of the J-statistic. The J-statistic, as explained earlier, will not be used in this study.

The results for the C-tests are given in Table 4.8. The abbreviations in the table are those explained at the beginning of this section. The first element in each group of entries in the table is the estimate of the nesting parameter, $\theta$. The second element is the standard of error of $\hat{\theta}$ and the C-statistic appears last.

Due to the fact that the concepts used in the tests are true only asymptotically and they are applied here to a relatively small sample. size $(N=44)$, a relatively large probability of type I error ( $\alpha$ ) will be allowed. The selected level is $\alpha=0.025$ corresponding to a critical level of $Z_{\alpha / 2}=2.24$ for a two-tailed test. The two-tailed tests will be used because interest is on hypothesis tests rather than discrimination and because of the disjointed nature of the parameter spaces of the hypotheses under test.

Looking at the columns of Table 4.8 labeled NS/SR and NS/Q which

Table 4.8 C-Statistic Test Results ${ }^{\text {a }}$

| HYPOTHESIS TESTED | NS/SR | SR/NS | NS/Q | Q/NS |  |
| :--- | :--- | :--- | :--- | :--- | :--- |
| YEAR | 1960 | 0.4289 | 0.5711 | 0.3717 | 0.6284 |
|  |  | 0.2797 | 0.2797 | 0.2582 | 0.2582 |
|  |  | 1.5334 | 2.0418 | 1.4394 | 2.4335 |
|  | 1963 | 0.2614 | 0.7386 | 0.2475 | 0.7525 |
|  |  | 0.1800 | 0.1800 | 0.1555 | 0.1555 |
|  |  | 1.4521 | 4.1025 | 1.5916 | 4.8385 |
|  |  | 0.1149 | 0.8851 | 0.1015 | 0.8985 |
|  |  | 0.2502 | 0.2502 | 0.2000 | 0.2000 |
|  |  | 0.4593 | 3.5368 | 0.5073 | 4.4922 |

${ }^{a}$ Entries in the table are in groups of three. In each group, the first line gives the coefficient estimate for $\theta$ and the second line its standard error. The third line gives the t-ratio for $\theta$ which is referred to in the text as the C-statistic.
give the results of testing the nutrient non-substitution hypothesis against the polynomial formulations, it can be seen that the estimated C-statistic is lower than the critical value of 2.24 for all the three years. Even when the significance level is increased to $\alpha=0.05$, the non-substitution hypothesis cannot be rejected at the resulting critical level of 1.96 .

The estimate of the variance of $\sqrt[N]{ } \hat{\theta}_{0}$ is given by $\hat{N}_{0}^{2}\left(X^{\prime} X\right)^{-1}$ where $X=\left(X^{\top} \hat{\gamma}-\hat{\mu}\right)$ and $\hat{\sigma}_{0}^{2}$ is the consistent estimator of $\sigma_{0}^{2}$ in (4.8.1). This variance estimate for $\sqrt{N} \hat{\theta}_{0}$ can be shown to be asymptotically biased upwards -- see for example Davidson and MacKinnon [1981, p. 787]. The resulting t-statistic used in the C-test will therefore have a tendency of not rejecting the null hypothesis more often than the CP statistic. This may explain why the non-substitution hypothesis could not be rejected in all the three years studied. However, a look at the results for the tests in which the polynomial formulations were tested against the non-substitution hypothesis (columns labeled SR/NS and Q/NS), reveals that using a critical level of 2.24 leads to the rejection of the polynomial formulations in 5 out of 6 cases.

The quadratic formulation was rejected in all the three years whereas the square root could not be rejected only in 1960. When the significance level is increased from $2.5 \%$ to $5 \%$, the estimated C-statistics indicate that both the square root and the quadratic formulations should be rejected in all the three years.

The fact that the C-statistic has a tendency of not rejecting the tested hypothesis further erodes the credibility of the polynomials since, in this case, the indications are that they should both be rejected and at a rather high level of significance.

No conclusive statements can be made solely on the basis of the C-tests, especially in view of the fact that the small sample size properties of the statistic used are unknown. It is also apparent from these results that the C-statistic is unable to decisively break the deadlock which existed between the non-substitution and the square root models for 1960 when comparisons were based on RVE. However, when the significance level is increased to $5 \%$, the tie is broken and the edge goes to the nutrient non-substitution hypothesis.

In order to carry out the non-nested hypothesis tests on the basis of the Cox-Pesaran (CP) procedure, some auxiliary estimations must be done, in addition to estimating the pair of models contained in the hypothesis under test. The relations to be estimated and the desired residual sums of squares were presented in Section 3.5 and will therefore not be discussed here again.

Three different formulations were compared and, since the polynomials were not tested against each other, only four pairs of hypotheses were tested in each year. Using the abbreviations introduced earlier, the four hypotheses tested were NS/SR, SR/NS, NS/Q and Q/NS. The first two were used to test the non-substitution and square root hypotheses and the other two tested the non-substitution and the quadratic hypotheses.

Results for the auxiliary estimations for the pair NS and SR are reported in Table 4.9. The first two columns labeled NS and SR give the residual sums of squares under the two tested hypotheses, adjusted by the respective degrees of freedom. The resulting quantities were referred to in Section 3.5 as $\tilde{\sigma}_{0}^{2}$ and $\tilde{\sigma}_{1}^{2}$. They are the unbiased estimators of $\sigma_{0}^{2}$ and $\sigma_{1}^{2}$ and have synonymously been referred to as RVE. Also,

Table 4.9 Results for Auxiliary Estimations for the CP Testing Procedure - NS and SR

|  | HYPOTHESIS TESTED PARAMETER |  | $\frac{S R}{z_{1}^{2}}$ | NS/SR |  | SR/NS |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\mathrm{e}_{10}{ }^{\prime} \mathrm{e}_{10}$ | $\mathrm{e}_{100} \mathrm{e}_{100}$ | $e_{01}{ }^{\prime} e_{01}$ | $e_{011}{ }^{\prime} e_{011}$ |
| YEAR | 1960 | 0.001839 | 0.001915 | 0.007020 | 0.006343 | 0.019292 | 0.006113 |
|  | 1963 | 0.001410 | 0.001881 | 0.034108 | 0.020557 | 0.025473 | 0.017871 |
|  | 1965 | 0.001850 | 0.002389 | 0.023534 | 0.013526 | 0.027014 | 0.014834 |

in earlier sections, these estimates were reported for variables measured in quintals/ha. As explained in Appendix Table A.1, all the response models were estimated with corn yield data measured in 100 quintals/ha and the coefficient estimates had to be retransformed back to quintals/ha. This is not done for the hypothesis test results because the basic quantities of interest are the residual sums of squares which are in comparable units for any pair of hypotheses under test. Hence, $\tilde{\sigma}_{0}^{2}$ appears in Table 4.9 as 0.001838 instead of 18.38 when measurements are retransformed back to quintals/ha.

The last four columns in Table 4.9 report the residual sums of squares from the auxiliary estimations mentioned above. They are derived from equations (3.40), (3.42), (3.43) and (3.44) of Section 3.5. Equations (3.40) and (3.44) were estimated by use of the multiple regression package provided in the TSP manual and therefore presented no computational difficulties.

Equation (3.43) required the MINOS mathematical programming approach similar to that used to estimate the original non-substitution model. It differed from the latter in the sense that the actual corn yield data in the RHS of the original problem was now replaced by corn yield data as estimated by the square root model. The two problems, however, closely resembled each other enough to enable starting the auxiliary estimation from a basis saved from the original problem. The auxiliary problem could also be started from any of the bases saved during the search for the optimal knot as previously discussed.

A not so preferred approach was a "cold start" because the problems were often more expensive to run when started without an initial solution. But, at times, it happened that a problem given a "warm start" wound up
in an infeasible pocket or failed because further iterations were unable to make improvements on the objective value, yet no difficulties were experienced when the same problem was started without an old basis to provide an initial solution.

There were also other options available for cases when a problem led to infeasibilities or was not converging. Such options involved rescaling of variables especially the RHS vector and changing of certain tolerances but no elaboration of these techniques will be undertaken here since there was no specific rule to go by.

Most of the computational difficulties were, however, encountered in the estimation of equation (3.42) in which the residual vector $\mathrm{e}_{10}$ was used as the RHS vector in the MINOS program. Since the vector $\mathrm{e}_{10}$ had some negative elements (unacceptable as the RHS elements in the mathematical program), a transformation was necessary before introducing $\mathrm{e}_{10}$ in the RHS. The problems encountered were related to variable scaling. At times the problem went infeasible or nonconverging simply because the variables were not scaled appropriately.

The scaling problem with regard to the estimation of equation (3.42) is discussed in detail towards the end of Appendix Table A.1. Briefly, it involved multiplying $e_{10}$ (which by the way, is the estimated vector. of residuals from equation (3.40)) by an appropriately chosen value (M) and adding a constant term $A$ such that all the elements of $e_{10}^{\star}=M e{ }_{10}+A$ were in the range $[0,15]$. The desired vector $e_{100}{ }^{\prime} e_{100}$ was then obtained as $\mathrm{e}_{100}^{\star} \mathrm{e}_{100}^{\star} / \mathrm{M}^{2}$ where $\mathrm{e}_{100}^{\star}$ is the estimated residual vector from the transformed model. Since the models were not forced through the origin, the additive factor (A) had no effect on the residuals and was therefore not involved in the retransformation back to the original
quantities. The objective was to choose a combination of $M$ and $A$ which led to $e_{10}^{\star}$ approximating the RHS vector of an already successful problem. The basis from the latter could then be used to provide an initial solution.

After a few trials with different combinations of $M$ and $A$, it was often possible to construct an optimally converging problem. The pairs ( $M, A$ ) actually used were as follows: $(100,2.5)$ for both NS/SR and NS/Q in 1960; $(10,2.5)$ for NS/SR and $(30,4)$ for NS/Q in 1963 and $(30,3)$ for both NS/SR and NS/Q in 1965.

Results for the auxiliary estimations for the hypothesis tests involving the non-substitution and the quadratic models (NS and Q) are presented in Table 4.10. The procedures for deriving the reported quantities were identical to those already discussed for NS and SR hence the table does not require additional explanations.

The results in Tables 4.9 and 4.10 were used to derive the CP statistics employing the formulae given in the part of Section 3.5 dealing with the CP testing procedure. The CP statistics are presented in Table 4.11 for the four hypotheses tested in each of the three years examined.

As stated earlier, the CP statistic is asymptotically distributed as $N(0,1)$ under the null hypothesis and hence the standard normal table provides the critical (rejection) points. Large negative values of the estimated statistic suggest that the null hypothesis ( $H_{0}$ ) should be rejected in favor of the alternative $\left(H_{1}\right)$. Large positive values on the other hand suggest that the null hypothesis should be rejected but in favor of a third hypothesis $\left(\mathrm{H}_{2}\right)$ differing from $H_{0}$ in some sense opposite to that in which $H_{1}$ differs from $H_{0}$. Once again, since interest is on hypothesis tests as opposed to discrimination among a set of

Table 4.10 Results for Auxiliary Estimations for the CP Testing Procedure - NS and Q

|  | HYPOTHESIS TESTED PARAMETER |  |  | NS/Q |  | Q/NS |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $e_{10}{ }^{\prime} \mathrm{e}_{10}$ | $e_{100}{ }^{\prime} e_{100}$ | $e_{01}{ }^{\prime} e_{01}$ | $e_{011}{ }^{\prime} e_{011}$ |
| YEAR | 1960 | 0.001839 | 0.002000 | 0.011458 | 0.010354 | 0.020897 | 0.009069 |
|  | 1963 | 0.001410 | 0.002071 | 0.044760 | 0.026251 | 0.028024 | 0.019450 |
|  | 1965 | 0.001850 | 0.002722 | 0.038240 | 0.023559 | 0.029478 | 0.016145 |

Table 4.11 The Estimated CP Test Statistics ${ }^{\text {a }}$

| HYPOTHESIS TESTED |  |  | NS/SR | SR/NS |
| :---: | :---: | :---: | :---: | :---: |
| YEAR | 1960 | -0.553 | -3.729 | -0.515 |
|  |  | $(-0.541)^{b}$ | $(-3.305)$ | $(-0.502)$ |

a The entries in the table are for $\bar{D}_{j}=T_{i} /\left[V_{i}\left(T_{j}\right)\right]^{\frac{1}{2}}$ where the subscript $i=0$ refers to the hypothesis being tested $\left(H_{0}\right)$ and $i=1$ to the alternative $\left(H_{1}\right)$. Testing $H_{0}$ against $H_{1}$ is summarized as $H_{0} / H_{1}$.
${ }^{b}$ Figures in parenthesis are the linearized CP statistics.
specifications, two-tailed tests will be employed.
When a critical value of 2.24 corresponding to a $2.5 \%$ significance level is used, the estimated CP statistics in Table 4.11 suggest that both polynomial specifications must be summarily rejected in favor of the non-substitution hypothesis. In fact, the figures for $S R$ and $Q$ tested against NS are such that the polynomial approximations cannot be accepted at any reasonable level of probability for a type I error.

Table 4.11 also reports the estimated linearized CP statistics (figures in parenthesis). This statistic is supposed to be more conservative at rejecting the tested hypothesis when the alternative is performing better than expected. This can be seen from the fact that the reported linearized CP statistics are slightly smaller than the corresponding CP statistics. The absolute values of the latter for SR/NS and Q/NS were, however, such that the linearization procedure still left the polynomial hypotheses in the rejection region.

The CP test results confirm those based on the crude C-statistic discussed earlier. The only exception was the 1960 square root form not being rejected by the C-statistic at the $2.5 \%$ significance level. The overall picture, however, was that there were no inconsistencies between the results obtained from the C-tests and those from the CP tests.

The C-tests are relatively easy to perform and if the correct asymptotic standard error for the estimated nesting parameter $(\hat{\theta})$ is available, may be sufficient as a basis for testing non-nested hypotheses. Judging from the results discussed above, it does not appear as if the power of the C-statistic was significantly diminished by the fact that use was made of a t-statistic employing an asymptotically biased estimate for the standard error for $\hat{\theta}$. The inability of the statistic to
reject the non-substitution hypothesis does not appear to be due to a problem inherent in the statistic. It may, however, be due to the inability of the polynomial hypotheses to reject the non-substitution hypothesis for the given sets of data. The C-tests may therefore be preferred to the more involved CP procedure if its power can be ascertained.

The CP statistic is invariant to the relative performance measures such as RVE and $\mathrm{R}^{2}$ for the hypotheses under test. This attribute was demonstrated here by the fact that it was able to reject the 1960 square root specification whereas the C-statistic and the individual RVE's indicated no clear superiority of the non-substitution model.

No definite statements can be made about the power (the probability of a statistic failing to reject a false hypothesis) of any of the tests discussed above especially for small samples. However, asymptotically, the CP statistic will reject the tested hypothesis with a probability of one whenever the alternative is true. The same cannot be said of the C-statistic in the form used, that is, without adjusting the variance of the estimated nesting parameter, $\hat{\theta}$. What can be said though is that for the small samples analyzed here, the C-statistic seems to have a performance similar to the CP statistic and is therefore worthy of consideration in non-nested hypothesis tests.

When the sample size was increased from 44 to 88 the two tests gave conflicting but otherwise not unexpected results. The data analyzed for the larger samples represented the years 1960-1961 and 1964-1965. The pooling of these two groups of years did not seem to pose such a serious risk with regard to the violation of the assumption of a block-wise constant variance. In other words, the two sets of data are those for which the disparity of the estimated standard errors of the regression between
any pair of blocks (years) is minimized. In spite of the discretion exercized in the selection of the years to be pooled together, there is no guarantee that the coefficient estimates reported in Tables 4.12 and 4.13 are maximum likelihood estimators. This fact does not, however, invalidate the tests based on these estimates since the hypotheses under test are given an equal treatment with respect to the included variables and data transformation.

The quadratic formulation, as can be judged from the tests based on the single-year models, performed poorer than both the non-substitution and square root forms. The two-year tests are therefore based only on the latter pair.

The entire solution to the non-substitution models are not reproduced in the appendix as was done for the single-year models on account of the fact that the former are too lengthy. The estimated yield maxima and knots are reported in Table 4.13 and the test results in Table 4.14. The pair of constants used to transform $e_{10}$ in order to estimate $e_{100}$ for 1960-1961 and 1964-1965 were (30,3) and (30, 2), respectively.

On the basis of the residual variance estimates (RVE), the square root specification performed better than the non-substitution model in both sets of data analyzed. Tests based on the CP statistic on the other hand indicate that the polynomial approximation should be rejected, at the $2.5 \%$ significance level, in favor of the non-substitution hypothesis in 1960-1961. For the same set of data, the C-statistic indicates that both hypotheses should not be rejected thus conflicting the test results based on the CP statistic by not rejecting the square root polynomial.

For 1964-1965 the CP statistic indicates, again, that the square root form should be rejected. This time, however, even the

Table 4.12 Results for the Estimation of the Square Root
Function-Pooled Data (quintals/ha)

| YEARS | 1960-1961 |  | 1964-1965 |  |
| :---: | :---: | :---: | :---: | :---: |
| COEFF. $\mathrm{S}_{1}$ | 52.1773 | $(7.5)^{\text {a }}$ | 47.6421 | (6.5) |
| $\mathrm{S}_{2}$ | 41.9599 | (6.1) | 45.4242 | (6.1) |
| $\mathrm{S}_{3}$ | -- |  | -- |  |
| $\mathrm{S}_{4}$ | -- |  | -- |  |
| PT | -1.4600 | (0.9) | -4.7615 | (3.0) |
| PTSQ ${ }^{\text {b }}$ | 7.5151 | (1.3) | 19.3350 | (3.3) |
| PTKT ${ }^{\text {b }}$ | 0.0408 | (0.04) | 0.5880 | (0.5) |
| KTSQ ${ }^{\text {b }}$ | 34.0637 | (6.1) | 25.1622 | (4.0) |
| KT | -7.3629 | (5.8) | -5.9548 | (4.2) |
| $\mathrm{R}^{2}$ | 0.7281 |  | 0.4643 |  |
| RVE | 26.3933 |  | 33.7762 |  |

${ }^{\text {a }}$ Entries in parenthesis are the coefficient t-ratios.
${ }^{\mathrm{b}}$ The abbreviations PTSQ, PTKT and KTSQ stand for (PT) ${ }^{\frac{1}{2}}$, (PTKT) $)^{\frac{1}{2}}$ and $(K T)^{\frac{1}{2}}$, respectively.


## Yield Maximum

102.9900
101.1000
91.7200
93.6900

KNOT
1.5500
1.6500

Potassium

| Block dummy variable | ZKD1 | 0.69785 | 0.72688 |
| :--- | :--- | :--- | :--- |
|  | ZKD2 | 0.61298 | 0.71834 |

Relative yield function
slope for KT $=$ [0-KNOT]
0.18321
0.12094

Change in slope at KNOT
$-0.18321$
$-0.12094$
Phosphorus
Block dummy variable ZPD1
$0.84371 \quad 0.70791$
ZPD2
1.20000
1.20000

Relative yield function
slope for PT $=[0-$ KNOT $]$
0.09421
0.20956

Change in slope at KNOT
-0.09421
-0.20956
RESIDUAL VARIANCE ESTIMATE
26.39332
33.77616

Table 4.14 Hypothesis Test Results for the Two-Year Models: NS and SR ( $N=88$ )

| HYPOTHESIS TESTED | NS | SR | NS/SR |  | SR/NS |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| PARAMETER | $\mathrm{O}_{0}^{2}$ | $\sigma_{1}^{2}$ | $e_{10}{ }^{\prime} \mathrm{e}_{10} 0$ | $e_{100}{ }^{\prime} e_{100}$ | $e_{01}{ }^{\prime} e_{01}$ | $e_{011}{ }^{\prime} e_{011}$ |


| $1960-1961$ | 0.00264 | 0.00257 | 0.02081 | 0.01272 | 0.23377 |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Aux. for CP tests |  | -2.14619 | -27.35131 |  |  |
| CP statistic |  | -2.02985 | -19.81916 |  |  |
| Linearized CP statistic |  | 2.06385 | 1.70118 |  |  |
| C-statistic |  |  |  |  |  |
| 1964-1965 | 0.00338 | 0.00295 | 0.02142 | 0.00945 | 2.11264 |
| Aux. for CP tests |  | -5.80015 | 0.01769 |  |  |
| CP statistic |  | -5.24266 | -340.92579 |  |  |
| Linearized CP statistic |  | 3.40375 | -143.60552 |  |  |
| C-statistic |  |  | 0.42791 |  |  |

non-substitution model must be rejected although not at levels of significance as high as those for the alternative hypothesis. The C-statistic also suggests rejection of the non-substitution model but not of the square root specification.

As can be seen from the summary of the test results given in Table 4.15, the nutrient non-substitution model is rejected by both the C-statistic in only one instance. The square root form on the other hand is rejected in all cases but for the pooled models in which it is not rejected by the C-statistic. This statistic, however, has a bias in favor of the tested hypothesis as the sample size gets larger for reasons already mentioned.

The linearized CP statistic ( $\bar{D} L$ ) was not able to reverse any of the conclusions made on the basis of the two statistics discussed above. For borderline cases, for example at the $2.5 \%$ significance level, the differences between $\overline{\bar{D}}$ and $\overline{\mathrm{D}} \mathrm{L}$ are too small to bring $\overline{\mathrm{D}}$ below the rejection points. Those differences are only large when the values of $\bar{D}$ are large, in which case, the linearized statistic does not even come close to the tabulated critical points as can be seen from Table 4.14.

The lopsided nature of the outcome of the hypothesis tests discussed above is hardly surprising in view of the fact that most crop response surfaces display significant plateaus when viewed over the major nutrients. The polynomials, with their intrinsic symmetric nature, are incapable of accounting for such plateaus or any sharp bends on the response surface.

For the data analyzed in this study, the presence of plateaus was evidenced not only by the fact that the non-substitution hypothesis, which explicitly accounted for them, could not be rejected in most of

Table 4.15 Statistical Status of Hypothesis Test Results Based on Different Testing Procedures ${ }^{\text {a }}$

| HYPOTHESIS TESTED | NS/SR | SR/NS | NS/Q | Q/NS |
| :---: | :---: | :---: | :---: | :---: |
| YEAR 1960 | $N R^{\text {b }}$ | R | NR | R |
|  | NR | R | NR | R |
|  | NR | NR* | NR | R |
| 1963 | NR | R | NR | R |
|  | NR | R | NR | R |
|  | NR | R | NR | R |
| 1965 | NR | R | NR | R |
|  | NR | R | NR | R |
|  | NR | R | NR | R |
| 1960-1961 | NR | R | -- | -- |
|  | NR | R | -- | -- |
|  | NR | NR | -- | -- |
| 1964-1965 | R | R | -- | -- |
|  | R | R | -- | -- |
|  | R | NR | -- | -- |

aThe symbols $R$ and NR stand for hypothesis "rejected" and "not rejected," respectively. Two-tailed tests were used at the $2.5 \%$ significance level and an asterisk indicates hypothesis rejected at 5\% significance level.
${ }^{\mathrm{b}}$ The entries are in groups of three referring to the three test statistics: CP, linearized CP and C-statistic.
the cases, but also by the fact that the quadratic hypothesis was rejected at significance levels higher than those of the square root hypothesis. The square root specification has a flatter surface than the quadratic form as was demonstrated in sub-section 4.2.1.

The plateau shape of the corn response data is by no means the overriding factor contributing to the rejection of the polynomial specifications. In other words, the non-substitution model is not to be accepted merely because it accounted for the plateau and hence fitted the data better than the alternative specifications. Indeed, as demonstrated by the 1960, 1960-1961 and 1964-1965 tests, the square root form which fitted the data just as well or better than the non-substitution hypothesis but had to be rejected by the CP statistic, a "good fit" alone is not a sufficient condition to guarantee acceptance of a given hypothesis.

The nutrient non-substitution model differs from the polynomial approximations in respects more basic than just accounting for the plateau surface. The most important of these differences is that the nutrient non-substitution model, as the name suggests, does not admit substitution among the major nutrients. The fact that the major nutrients do not substitute each other in plant nutrition is now widely acknowledged. It is also true that the results of this study indicate that the nutrient non-substitution hypothesis performs better than the alternative polynomial approximations for the given data sets. This, however, does not mean that the hypothesis of nutrient non-substitutability is unequivocally sustained. The model was not intended for proving laws of plant nutrition. It is nevertheless based on such laws. To that extent, the fact that the nutrient non-substitution hypothesis rejects polynomial hypotheses (which violate some conventional biological theories) and is in turn
rejected by the latter in only one instance, accords credibility to the model based on nutrient non-substitution. The question of "economic substitution" deriving from applied nutrients does not arise because the models were based on total nutrient supplies.

In the absence of alternative models with assumptions less damaging to established biological principles than those implicit in the nutrient non-substitution model, advocating the use of the latter in crop response analysis would be considered as scientifically sound. The polynomial approximations do not seem to be supported by the hypothesis test results of this study, and even on the basis of measures of relative fit such as RVE the non-substitution model was highly competitive.

Chapter 5: Conclusions and Directions for Future Research

In this study, the model of corn response was formulated as a simultaneous equations system with response ( $y$ ) and soil test ( $b^{*}$ ) as the endogenous variables. The system was, however, assumed to be recursive hence the two equations were estimated separately as opposed to applying simultaneous equations techniques.

The equation estimating $b^{*}$ was referred to as the carryover function. The dynamic aspects of the fertilization problem were accounted for by introducing lagged values of $y$ and $b$ * exogenously in the carryover function. The carryover function provided estimates of conversion factors ( $\lambda$ ) used to calibrate the soil test values in units of applied fertilizer.

The basic time framework in the carryover function was a rotation consisting of four years. Hence, the estimated $\lambda$-values could not be used directly in the crop response function based on seasonal (one year) data. The seasonal $\lambda$-values were extrapolated from four-year and twoyear $\lambda$-values as estimated by the carryover function. The extrapolated values were $12.3 \mathrm{~kg} / \mathrm{ha}$ and $1.6 \mathrm{~kg} / \mathrm{ha}$ for P and K , respectively. These estimates were then used, together with data on applied fertilizer, to construct total seasonal availabilities of the two nutrients P and K used as the independent variables in the corn response function. In future, loss of efficiency in coefficient estimates at both stages of estimation may be avoided by devising a formulation capable of providing the desired seasonal $\lambda$-values directly. There is little complication if soil sampling is done for each plot every season. This procedure can, however, prove to be financially infeasible. Emphasis must therefore be directed to specifications and estimation of carryover functions based on data
involving more than one crop in a rotation and scarce soil test information.

The appropriateness of two kinds of specification of corn response functions was examined. The first kind was a polynomial approximation which has frequently been applied by many agricultural economists and soil scientists and by now considered as the conventional form for depicting crop response surfaces. The second kind was a nutrient non-substitution formulation constrained to have a flat surface beginning at a total soil nutrient level corresponding to maximum crop yield. This nutrient level was obtained as the $k^{\text {th }}$ knot in a spline function with $k+1$ knots.

The non-substitution model generalizes Liebig's law of the minimum and has flexibility as one of its virtues. The expected corn yield, in the two nutrient case examined in this study, is given by $\hat{A} \operatorname{Min}\left[f_{p}(P T), f_{k}(K T)\right]$ where $\hat{A}$ is the estimated yield maximum and the arguments in $f_{p}(\cdot)$ and $f_{k}(\cdot)$ representing estimated total availability of $P$ and $K$, respectively.

The individual response functions $f_{p}(\cdot)$ and $f_{k}(\cdot)$ can be approximated by any appropriate form including, or course, polynomials or any other function allowing yield depression. However, the computational aspects must be considered when determining the functional forms of $f_{p}(\cdot)$ and $f_{k}(\cdot)$. In this study, the individual response functions were approximated by linear splines and estimated by applying a nonlinear mathematical programming technique to a nonlinearly constrained problem.

The techniques used hitherto in fitting the law of the minimum involved sorting out data so as to separate those in which either PT or KT was limiting growth, a procedure which is not only cumbersome but also
statistically questionable. The mathematical approach adopted here facilitated simultaneous estimation of $f_{p}(\cdot)$ and $f_{k}(\cdot)$ and required no prior sorting of data or special experimental designs.

The computational problems encountered were those not uncommon in iterative estimation procedures which typically related to convergence of the iterations to a stable point. It was found that the iterations converged fairly quickly (less than 4 major iterations) whenever an initial solution was provided. However, the effect of different starting points and the general problems of stability were left for future research.

Other computational problems resulted from scaling of variables.
This was due to the fact that the MINOS/AUGMENTED program reacted differently to different scalings. The scaling also led to a loss in variability in data thus leading to convergence problems. It is recommended therefore that alternative formulations and estimation of the law of the minimum be sought. One feasible approach would be to make the constraint set as formulated here free of the nonlinearities. This can probably be accomplished by introducing the nonlinear constraints ( $\left.S_{k t} S_{p t} \geq 0\right)$ in the criterion function as penalty functions thereby leaving only the linear spline functions in the constraint set.

Two polynomial formulations of the response of corn were estimated: the square root and the quadratic forms. The "fit" as judged by the $R^{2}$ was low (in the neighborhood of $55 \%$ ) but could have been improved by including more sources of variation in yield. Both forms indicated lack of interaction between P and K. Generally, there was little response to $P$ pointing to an oversupply of this nutrient in the experimental plots.

The results based on the polynomial formulations were verified by
those of the nutrient non-substitution model which indicated that the maximum corn yield could be obtained by total nutrient supplies no more than $165 \mathrm{~kg} /$ ha for both P and K . It is possible, however, that this value slightly overestimated the level of $P$ required for maximum corn yield in most of the years examined. Had different knots been used for the spline functions of $P$ and $K$, it is likely that the optimal knots for $P$ would have been lower than those reported in subsection 4.2.2 and hence the accompanying residual sums of squares.

The decision to adopt equal knots for the two spline functions was on the grounds of computational convenience and admittedly sacrificed a certain amount of efficiency. In fact, the very idea of using linear splines to approximate the individual response functions was a simplification which may have been traded for some degree of precision. These are certainly aspects which could benefit from further investigation and professional exchange of ideas.

From the foregoing discussion, it should be apparent that the nonsubstitution model is still at a testing stage. The non-nested hypothesis tests carried out in this study showed that the model, even in a simplistic (linear and plateau) form, generally performed better than the polynomial approximations.

The three different test statistics gave consistent results for small samples ( $N=44$ ). When $N$ was increased to 88 the $C$-statistic had results conflicting those of the Cox-Pesaran statistic.

The C-statistic is not reliable asymptotically and was used only as a preliminary test to be compared with the more involved Cox-Pesaran statistic.

On the basis of the Cox-Pesaran statistic, it can be concluded that
the two polynomial forms could not be supported against the nutrient nonsubstitution hypothesis given the data analyzed in this study. The linearized Cox-Pesaran statistic was not found useful since it could not reverse decisions made on the basis of the unlinearized statistic.

It is worth emphasizing that the samples analyzed here may not be considered as large enough to warrant application of asymptotic concepts. Further tests with larger samples may have to be conducted before making positive statements about the appropriateness of the non-substitution hypothesis vis-à-vis the polynomial approximations or any other specification for that matter. The results of this study should not be used as a basis for discouraging the application of polynomial forms in crop response analysis. The non-substitution model is proposed as a flexible alternative, particularly in cases where the mathematical form of the conventional approximations do not accurately represent the underlying biological phenomena or the physical configuration of the response surface.

As indicated earlier in this study, the non-substitution model should be accepted not merely because it fits the data better than the polynomials. The hypothesis tests showed that this is not a sufficient condition to guarantee acceptance of a given specification. The biases tested for in this study were more fundamental than differences based on measures of relative (statistical or economic) performance. It is hoped that the model will achieve professional acceptance as a generalized tool which is not only biologically more appealing but is also statistically more appropriate than the conventional approximations currently employed in crop response analysis.

## Appendix

## A. 1 Data Transformations

Corn, soybean, and wheat data were originally reported in bushels (bu) per acre. Hay was reported in 1b/ac. All these yield data were consequently converted to $\mathrm{kg} / \mathrm{ha}$ using the following unit conversion factors: $11 \mathrm{~b}=0.454 \mathrm{~kg}, 1$ ha $=2.471 \mathrm{ac} ., 1 \mathrm{bu}$ of soybean and wheat $=27.2 \mathrm{~kg}$, and 1 bu of corn (shelled) $=25.4 \mathrm{~kg}$. To facilitate easy computation, the yield data were further divided by 100.

The soil testing procedures varied over the years. From 1954 to 1964 the Purdue Soil Testing Laboratory reported the phosphorus soil test results in pounds of $\mathrm{P}_{2} \mathrm{O}_{5}$ per 2 million pounds of soil. Test results for potassium were similarly reported in pounds of $\mathrm{K}_{2} 0$ per 2 million pounds of soil. From 1965 to 1967 , both P and K test results were reported on an elemental basis. The extraction of $P$ and $K$ was done by shaking 5 g of soil with 15 ml of 0.7 HCl in a shaker for 2 minutes. In the case of P, this procedure was dropped in favor of Bray P1 in 1968.

The analysis carried out in this study used soil tests, $b^{\star}$, measured in $\mu \mathrm{g} / \mathrm{g}$ and for phosphorus, units have been standardized to Bray. Pl.

Denote the 1954-1964 test results by $P_{1}$ and $K_{1}$ and the 1965-1967 test results by $P_{2}$ and $K_{2}$. The change in 1968 in the testing procedures affected only $P$ hence the test results from then on are denoted by $P_{3}$ and $K_{2}$. Given that $\mathrm{P}_{2} \mathrm{O}_{5} \times 0.4366=P$ and that $\mathrm{K}_{2} 0 \times 0.83=\mathrm{K}$, the original test results were converted to elemental units as follows:

$$
\begin{align*}
& P_{1} \times 0.4366=P_{2}  \tag{a.1}\\
& K_{1} \times 0.83=K_{2} \tag{a.2}
\end{align*}
$$

$\mathrm{P}_{2}$ and $\mathrm{K}_{2}$ are in lb (of P and K ) per 2 million lb of soil.
The Purdue extraction procedure used prior to 1968 extracted approximately twice the amount obtained by Bray P1. ${ }^{1 /}$ The final soil test values are therefore obtained as follows:

$$
\begin{align*}
& P_{2} / 2 \times 2=b_{p}^{\star}  \tag{a.3}\\
& \mathrm{P}_{3} / 2=b_{p}^{\star}  \tag{a.4}\\
& \mathrm{K}_{2} / 2=b_{k}^{\star}
\end{align*}
$$

Notice that (a.3) and (a.5) convert $P_{2}$ and $K_{2}$ which are in 1 b of $P$ and $K$ per 2 million 1 b of soil to lb of P and K per 1 million lb of soil or ppm. $P_{3}$ on the other hand is divided only by 2 since it is already in Bray P1 units.

The phosphorus applied to corn-1 (CP) and wheat (WP) and similarly the potassium applied to corn-1 (CK) and wheat (WK), all derived from Appendix Table A.2, were divided by 100. Thus both yield and application rates were measured in quintals/ha ( $100 \mathrm{~kg} / \mathrm{ha}$ ) while soil test values were measured in $\mu \mathrm{g} / \mathrm{g}$ and, in the case of P, in Bray Pl units.

The results for the estimation of the carryover functions given in Section 4.1 are, however, in $\mathrm{kg} / \mathrm{ha}$. The scaling affected only the. coefficients corresponding to the scaled variables, not their t-ratios. Since in the carryover function the variables CP, CK, WP, WK, and all the yield data were measured in quintals/ha, the estimates of $\lambda_{4}^{-1}, \lambda_{2}^{-1}$, and $\alpha$ were retransformed to $\mathrm{kg} / \mathrm{ha}$ by dividing $\hat{\lambda}_{4}^{-1}, \hat{\lambda}_{2}^{-1}$, and $\hat{\alpha}$ by 100. In the case of the proportionality constants, the reported values are their

[^7]reciprocals $\lambda_{4}$ and $\lambda_{2}$ which were used to extrapolate the desired estimates $\lambda_{p}$ and $\lambda_{k}$.

For the estimation of the response functions, the corn yield data (as given in Appendix Table A.3) were further divided by 100. The independent variables were, however, left in quintals/ha. The tables reported in Section 4.2 have coefficients already retransformed to quintals/ha by multiplying by 100 .

In carrying out the non-nested hypothesis tests using the CoxPesaran (CP) procedure, there was special interest in sums of squared residuals from the models under test. Since the variables in the polynomials and the nutrient non-substitution models were scaled to the same degree, no retransformation of the residuals or their sums of squares was required; i.e., results from the estimation of the models under test were directly comparable.

There was one important exception to the statement made above. The regression

$$
\begin{equation*}
\dot{\mu}+x^{\top} b_{0}+\varepsilon_{10} \tag{a.6}
\end{equation*}
$$

in which the left-hand side vector represents estimates of the expected yield from the non-substitution model, gave the residuals $e_{10}$ used as the right-hand side (RHS) vector in the non-substitution model as . explained in the text. The magnitude and signs of some of the elements in the vector $\mathrm{e}_{10}$ made them unacceptable as members of the RHS in the mathematical programming setup which provided the vector $e_{001}$. The model estimated was therefore

$$
\begin{equation*}
\mathrm{e}_{10}^{\star}=\mu+\varepsilon_{100} \tag{a.7}
\end{equation*}
$$

given the constraints mentioned in the text and where $e_{10}^{\star}=M e_{10}+A$.
There is no rigorous rule concerning the choice of the terms $M$ and $A$
as long as the indicated manipulation leaves all the elements of $\mathrm{e}_{10}^{\star}$ non-negative.

The MINOS program used for estimating (a.7) is sensitive to variable scaling hence more than one combination of $M$ and $A$ may have to be tried out before obtaining a suitable vector $\mathrm{e}_{10}^{*}$. Improper scaling may lead to non-converging problems or infeasibilities. In this particular study, success was achieved by scaling all variables so as to make their values lie between 0.0 and 15.0 . The corn-1 yield data $(Y)$, after the transformation already mentioned in this appendix, had values less than 2.0. It therefore turned out that after estimating the non-substitution model with the corn-1 yield data as the right hand side (RHS) in the MPS file of the MINOS program, convergence of problems in which only the RHS of the MPS file was changed (as in the case of (a.7)) was achieved much faster if the new RHS vector (in this case $e_{10}^{*}$ ) was approximately of the same magnitude as the outgoing RHS vector (Y). The aim, therefore was to choose values of $M$ and $A$ which led to a vector $\mathrm{e}_{10}^{*}$ which approximately maintained the identity of the original problem with $Y$ as the RHS. Once this was achieved, the problem in (a.7) was then given a "warm start" from the basis saved from the original problem with $Y$ as the RHS vector. This procedure led to a significant reduction in the computer processing time and the number of major iterations required before convergence was achieved.

All the data fitting problems estimated here using the MINOS program involved functions not forced through the origin. The additive factor A therefore merely shifted the elements of $e_{10}$ away from the negative quadrant. The desired residual vector was therefore obtained as

$$
\begin{equation*}
e_{100}=e_{100}^{\star} / M \tag{a.8}
\end{equation*}
$$

where $e_{100}^{\star}$ is the vector of residuals obtained from the estimation of (a.7). $e_{100}{ }^{\prime} e_{100}$ which was needed in testing the nutrient non-substitution hypothesis was then obtained as $e_{100}^{*}{ }^{\prime} e_{100} / M^{2}$. The retransformation involves only the scalar $M$ and not the additive factor $A$.

## A. 2 Application Rates for $P$ and $K$ on <br> Purdue Fertility Experiment (kg/ha)

Treatment

| \# | Before Corn-1 @ 4 Yrs. | Corn-1 | Soybean | Wheat | Hay/Corn-2 |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | P-K | P-K |  | P-K |  |
| 1 | $0-0$ | 4.84-9.20 | -- | 14.51-27.58 | -- |
| 2 | $0-183.89$ | $0-0$ | -- | $0-0$ | -- |
| 3 | $0-183.89$ | 4.84-0 | -- | 14.51-0 | -- |
| 4 | $0-183.89$ | 12.09-0 | -- | 36.27-0 | -- |
| 5 | $0-183.89$ | 24.18-0 | -- | 72.55-0 | -- |
| 6 | $0-551.67^{\text {a }}$ | $0-0$ | -- | $0-0$ | -- |
| 7 | $0 \quad-551.67$ | 4.84-0 | -- | 14.51-0 | -- |
| 8 | $0 \quad-551.67$ | 12.09-0 | -- | 36.27-0 | -- |
| 9 | 96.73-183.89 | 0-0 | -- | $0-0$ | -- |
| $10^{\text {b }}$ | 77.38-183.89 | 4.84-0 | -- | 14.51-0 | -- |
| 11 | 48.37-183.89 | 12.09-0 | -- | 36.27-0 | -- |
| 12 | 96.73-551.67 | $0-0$ | -- | $0-0$ | -- |
| $13^{\text {b }}$ | 77.38-551.67 | 4.84-0 | -- | 14.51-0 | -- |
| 14 | 48.37-551.67 | 12.09-0 | -- | 36.27-0 | -- |
| 15 | 193.46-0 | 4.84-0 | -- | 14.51-0 | -- |
| 16 | 193.46- 0 | 4.84-22.98 | -- | 14.51-68.96 | -- |
| 17 | 193.46-183.89 | $0-0$ | -- | $0-0$ | -- |
| $18^{\text {b }}$ | 193.46-183.89 | 4.84-0 | -- | 14.51-0 | -- |
| 19 | 193.46-183.89 | 4.84-22.98 | -- | 14.51-68.96 | -- |
| 20 | 193.46-551.67 | $0-0$ | -- | $0-0$ | -- |
| $21^{\text {b }}$ | 193.46-551.67 | 4.84-0 | -- | 14.51-0 | -- |
| 22 | 193.46-551.67 | 12.09-0 | -- | 36.27-0 | -- |

${ }^{\text {a }}$ Changed from $367.78 \mathrm{~kg} / \mathrm{ha}$ to $551.67 \mathrm{~kg} / \mathrm{ha}$ in 1957.
${ }^{\mathrm{b}}$ Four treatments on which application of $P$ was stopped in 1973.
A. 3 Corn-1 yield (quintals/ha), soil tests (SP \& SK in $\mu \mathrm{g} / \mathrm{g}$ ) and applied fertilizer (CP \& CK in quintals/ha)

| 1400 | CUKN-1 | SP | Cr | U1 | U2 | U3 | 04 | 1461 | CUKN-1 | SP | CP | DI | 02 | DS | D4 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | OU. Y 1 | 9.06 | 0.11 | 1 | 0 | 0 | 0 | 1 | 67.27 | 2.40 | 0.11 | 0 | 1 | 0 | 0 |
| 2 | H 4.66 | 7.04 | 0.00 | 1 | 0 | 0 | 0 | 2 | 81.46 | 4.04 | 0.00 | 0 | 11 | 0 | 0 |
| 3 | 94.18 | 7.42 | U. 11 | 1 | 0 | 0 | 0 | 3 | 84.32 | 3.71 | 0.11 | 0 | 1 | 0 | 0 |
| 4 | 100,38 | 10.04 | 0.28 | 1 | 0 | 0 | 0 | 4 | 64.72 | 12.88 | 0.28 | 0 | 1 | 0 | 0 |
| 5 | 90.84 | 15.28 | 0.55 | 1 | $\checkmark$ | 0 | 0 | 5 | 86.74 | 8.95 | 0.55 | 0 | $\Gamma$ | 0 | 0 |
| 6 | 100.75 | 7.97 | 0.00 | 1 | 0 | 0 | 0 | 6 | 79.48 | 3.27 | 0.00 | 0 | 1 | 0 | 0 |
| 7 | 94.76 | 13.10 | 0.11 | 1 | 0 | 0 | $U$ | 7 | 66.06 | 7.53 | 0.11 | 0 | 1 | 0 | 0 |
| 8 | 94.49 | 17.14 | 0.28 | $\Gamma$ | 0 | 0 | 0 | 8 | 91.51 | 6.00 | 0.28 | 0 | 1 | 0 | 0 |
| $y$ | 40.58 | 10.04 | 2.20 | 1 | 0 | 0 | 0 | 9 | 90.21 | 5.89 | 2.20 | 0 | 1 | 0 | 0 |
| 14 | Ho.b0 | 10.70 | 1.07 | 1 | 0 | 0 | $u$ | 10 | 93.25 | 6.22 | 1.87 | 0 | 1 | 0 | 0 |
| 11 | 92.38 | 10.48 | 1.38 | 1 | 0 | 0 | 0 | 11 | 83.51 | 7.86 | 1.38 | 0 | 1 | 0 | 0 |
| 12 | 90.52 | 18.77 | 2.20 | 1 | 0 | 0 | 0 | 12 | 07.79 | 4.80 | 2. 20 | 0 | 1 | 0 | 0 |
| 13 | 42.68 | 9.93 | 1.87 | 1 | 0 | 0 | 0 | 13 | 87.98 | 6.66 | 1.87 | 0 | 1 | 0 | 0 |
| 14 | 97.77 | 11.35 | 1.38 | 1 | 0 | 0 | 0 | 14 | 86.45 | 6.55 | 1.38 | 0 | 1 | 0 | $0^{-}$ |
| 15 | 19.30 | 20.20 | 4.51 | 1 | 0 | $\cup$ | 0 | 15 | 39.68 | 10.12 | 4.51 | 0 | 1 | 0 | 0 |
| 16 | 97.40 | 20.41 | 4.51 | 1 | 0 | 0 | 0 | 16 | 74.65 | 10.31 | 4.51 | 0 | 1 | 0 | 0 |
| 17 | 97.90 | 14.41 | 4.40 | 1 | 0 | 0 | 0 | 17 | 87.30 | 12.12 | 4.40 | 0 | 1 | 0 | 0 |
| 10 | 100.15 | 20.20 | 4.51 | 1 | 0 | 0 | 0 | 16 | 85.19 | 19.65 | 4.51 | 0 | 1 | 0 | 0 |
| 14 | 43.87 | 27.51 | 4.51 | 1 | 0 | 0 | 0 | 19 | 66. 61 | 22.48 | 4.51 | 0 | 1 | 0 | 0 |
| 20 | 103.29 | 11.02 | 4.46 | 1 | 0 | 0 | 0 | 20 | 86.99 | 15.94 | 4.40 | 0 | 1 | $\square$ | 0 |
| 21 | 101.43 | 24.23 | 4.51 | 1 | 0 | 0 | 0 | 21 | 90.00 | 13.45 | 4.51 | 0 | 1 | 0 | 0 |
| 22 | 93.19 | 20.02 | 4.68 | 1 | 0 | 0 | 0 | 22 | 91.08 | 16.59 | 4.68 | 0 | 1 | 0 | 0 |
| <3 | 85.25 | 3.71 | 0.11 | 1 | 0 | 0 | 0 | 23 | 72.79 | 7.31 | 0.11 | 0 | 1 | 0 | 0 |
| 24 | 44.40 | 1.42 | 0.00 | 1 | U | 0 | 0 | 24 | 84.75 | 6.11 | 0.00 | 0 | 1 | 0 | 0 |
| 23 | 47.61 | 4.00 | 0.11 | 1 | 0 | 0 | 0 | 25 | 84.44 | 10.39 | 0.11 | 0 | 1 | 0 | 0 |
| $\cdots$ | 95.79 | 6.81 | 0.28 | 1 | 0 | 0 | 0 | 26 | 84.15 | 10.04 | 0.25 | 0 | 1 | 0 | 6 |
| Cl | 105.91 | 0.75 | 0.55 | 1 | 0 | 0 | 0 | 27 | 01.70 | 10.15 | 0.55 | 0 | 1 | 0 | 0 |
| 20 | 40. 29 | 5.35 | 0.00 | 1 | 0 | 0 | 0 | 28 | 84.20 | 0.19 | 0.00 | 0 | 1 | 0 | 0 |
| 24 | 92.69 | 7.45 | 0.11 | 1 | 0 | 0 | 0 | 29 | 92.38 | 7.86 | 0.11 | 0 | 1 | 0 | 0 |
| 50 | 100.56 | 11.90 | 0.28 | 1 | 0 | 0 | 0 | 30 | 40.46 | 23.79 | 0.28 | 0 | 1 | 0 | 0 |
| 31 | 94.18 | 6.88 | 2.20 | 1 | 0 | 0 | 0 | 31 | 85.19 | 11.40 | 2.20 | 0 | 1 | 0 | 0 |
| 32 | 95. 11 | 11.68 | 7.67 | 1 | 0 | 0 | 0 | 32 | 86.12 | V.30 | 1.81 | 0 | T | 0 | 1 |
| 55 | 94. 59 | 4.01 | 1.38 | 1 | 0 | 0 | 0 | 33 | 66.12 | 12.12 | 1.38 | 0 | 1 | 0 | 0 |
| 34 | 104.55 | 6.71 | 2.20 | 1 | 0 | 0 | 0 | 34 | 6b, U4 | 10.70 | 2.20 | 0 | 1 | 0 | 0 |
| 35 | 105.04 | 8.19 | 1.87 | 1 | 0 | 0 | 0 | 35 | 86.74 | T2.22 | 1.87 | 0 | 1 | 0 | 0 |
| 30 | 99.01 | 11.46 | 1.38 | 1 | 0 | 0 | 0 | 36 | 69.71 | 11.02 | 1.30 | 0 | 1 | 0 | 0 |
| 37 | 80.49 | 26.16 | 4.51 | 1 | 0 | 0 | 0 | 37 | 80.66 | 23.79 | 4.51 | 0 | 1 | 0 | 0 |
| 30 | IUU. 75 | 22.27 | 4.51 | 1 | 0 | 0 | 0 | 38 | 86.12 | 19.10 | 4.51 | 0 | 1 | 0 | $\checkmark$ |
| 34 | 95.60 | 21.28 | 4.40 | 1 | 0 | 0 | 0 | 39 | 67.73 | 16.15 | -. 40 | 0 | 1 | 0 | 0 |
| 40 | 101.00 | 14.97 | 4.51 | 1 | 0 | 0 | 0 | 40 | 88.04 | 16.48 | 0.51 | 0 | 1 | 0 | 0 |
| 41 | 105.34 | 21.07 | 4.51 | 1 | 0 | 0 | 0 | 41 | 88.97 | 18.77 | 4.51 | 0 | 1 | 0 | 0 |
| 42 | 97.15 | 20.06 | 4.40 | 1 | 0 | 0 | 0 | 42 | 91.45 | 16.92 | 4.40 | 0 | 1 | 0 | 0 |
| 45 | 104. 55 | 10.25 | 4.51 | 1 |  | 0 | 0 | 43 | H8. 16 | 17.46 | 4.51 | 0 | 1 | 0 | 0 |
| 44 | 101.95 | 20.96 | 4.68 | 1 | 0 | 0 | 0 | 44 | 89.40 | 20.41 | 4.65 | 0 | 1 | 0 | 0 |







$$
2
$$
















Y8


正
 m ~0000000000,000000000000000000005000000100010000


A. 4 Subroutine: CALCFG

| 100 | \$9t | LIthary | 00000100 |
| :---: | :---: | :---: | :---: |
| 200 |  | SUBHOUTINE CALCFG(MOUE, $\mathrm{N}, \mathrm{X}, \mathrm{F}, \mathrm{G}, \mathrm{NSTATE}$, NPROE) | 00000200 |
| 300 |  | LMPLICII KEAL (A-H:O-I) | 00000300 |
| 400 |  | REAL X(N),G(N) | 00000400 |
| 500 | C- |  | 00000500 |
| 600 |  | NTEN/S | 00000600 |
| 700 | C- |  | 00000700 |
| 800 |  | $F=0$. | 00000800 |
| 960 |  | $00500131 . N T$ | 00000900 |
| 1000 |  | $F=F \cdot X(2 \oplus N T+1) * X(2 ヵ N T+1)$ | 00001000 |
| 1100 |  | $G(I)=0$. | 00001100 |
| 1200 |  | G(ItMI)E0. | 00001200 |
| 1300 |  | $G(I+2 * N T)=x(1+2 * N T)$ | 00001300 |
| 1400 | 500 | Cuntinue | 00001400 |
| 1500 | C- |  | 00001500 |
| 1600 |  | FE F*, 5 | 00001600 |
| 1700 |  | RETURN | 00001700 |
| 1800 |  | ENLL | 00001800 |

## A. 5 Subroutine: CALCON



## A． 6 The MPS for the 1960 non－substitution problem



| 6700 | $E$ | WUmYOU1 |
| :---: | :---: | :---: |
| －800 | $E$ | MUNYUI？ |
| $4800^{\circ}$ | $E$ | ROWTOUS |
| $30 \cup 0$ | $E$ | MUnYOO4 |
| 5140 | $E$ | MUnTOOS |
| －52v0 | $E$ | सThTOO6 |
| 5300 | $E$ | HONYUOT |
| Savo | $E$ | mUnYU06 |
| 5500 | E | \％णヘT0 $\%$ |
| 5600 | $t$ | mumroio |
| 5740 | L | MUnY011 |
| $58 \cup 0$ | $E$ | nuntule |
| 3900 | $t$ | R（IwYU13 |
| －0vo | $E$ | HumYUia |
| －1vo | $t$ | kumruls |
| －2 20 | $E$ | k川mru16 |
| －3v0 | L | momrul7 |
| ＊4 U | $E$ | Mumrutz |
| －SuO | 8 | KUwruly |
| 6600 | $t$ | NUAYURO |
| 6700 | $E$ | kuñosi |
| －8u0 | $t$ | mimruez |
| －900 | t | mumrues |
| 1000 | $\varepsilon$ | MUnYU24 |
| 1100 | $E$ | wUnYu2S |
| 1200 | $E$ | mumyo 6 |
| 1300 | $t$ | WUnYOET |
| 1900 | $E$ | MUmYOes |
| $75 \cup 0$ | $E$ | MUnYu24 |
| 1600 | E | m0円Vu30 |
| 1700 | ＋ | mumyus |
| 1800 | $\varepsilon$ | WUWYO32 |
| 1900 | $E$ | kunyus3 |
| －uvo | $\bullet$ | numrus4 |
| 100 | E | m0¢ru35 |
| B200 | $E$ | NUMY030 |
| －300 | $E$ | MUWYUS7 |
| 840 | $E$ | WU⿴囗U36 |
| 8500 | $E$ | ROMYOS＊ |
| Bbub | e | NOMTU40 |
| 870 | $t$ | WUMYU41 |
| d 60 | $E$ | ROmYO4E |
| 6900 | $E$ | MOmYO43 |
| $90 \cup 0$ | $E$ | KUmYO4 |
| 9100 | $E$ | RO風2001 |
| 9200 | $E$ | NUW2K002 |












## A. 7 Solution for the 1960 non-substitution problem

## 

```
MINOSONOISK
```



```
    HOST IS (RSO314T)MINOS/MOST:
    BIND_CALCFG FROM_O甘JECT/USERII
    OINO CALCON FROM OBJECT/USER2:
```



```
    BEGIN AINDING CALCFG OF OUT FROM OBJECT/USER!
            CALCFG (02,NOO2) CHANGED TO (02,0053)
            7010 (01.0004) CHANGED TO (01.0004)
            <SEG DICT ITEMP - (01,0002) CHANGED 10 (01,000C) 03 0000022001EO
    END OF BINDING CALCFG
    AEGIN BINDING CALCON OF ,OUT FAOM OBJECT/USERZ
    -...CALCON (02,0002) CHANGED 10 (02,0055)
            7010 (01,0004) CHANGED 10 (01,000年)
            <SEG DICT |TEH> (01,0002) CHANGED TO (01,000E) = 03 0000018001FZ
END OF BINDING CALCON...............
```



```
NUMBER OF ERRORS OETECTED % 0
HOST FILE E (RSOSIATIMINOS/HOST ON DISK.
SEGMENT DICTIONARY LENGTH: IIA, GLOBAL STACK SIIE - 52G. STACK ESTIMATE E SIZ.
CORE ESIIMAIE E 2403I mORDS. CODE FILE LENGTH = 1859 DISK SEGMENTS.
BINDING IIME : 12 SECONDS ELAPSED, 2.29 SECONOS PROCESSOR, 20.39 SECONDS I/O.
```


sPECS FILE

BEGIN
NEWGASIS FILE II
OLD GASIS FILE 10
SOLUPIDN FILE 15
INPUT FILE 9
nows 700
COLUNNS 50@
ELEMENTS 1000
JACOUIAN SPARSE
majur iterations 100
nonlinear constralnts at
monlineah jacoulan variables 8
NONLINEAR OBJECIIYE_VARIABLES. 132
SUPERBASICS IS5
hessian dimension 130
LU ROM TOLERANCE 0.001
lu COL TOLERANCE 0.1
ERROR MESSAGE LIMIT 10
UPPER BOUND 3.0
END
(
MPS INPUT DATA.
 ..... 700
COLUMN LIHI! ..... 500
ELEMENTS LIMIT (COEFFS) ..... 3000
LISTLIMIT.................ERHOR MESSAGE LIMIT.PHANTOM ELEMEMIS........
FILES.
MPS FILE (INPUT FILE). ..... 9 ..... 15 ..... 0
0
OLD BASIS FILE (MAP)... NEW BASIS FlLE (MAP)!e! BACKUP BASIS FILE...... 
FRE QUENCIES.
LOG ITERATIONS........... ..... 1
SAYE NEH HASIS MAP...e日 ..... 100
. FACIORILE SIMYERI)CHECK RUM ERRORFEASIBILITY TOLERANCEALP PARAMEIERS.ITERATIOMS LIBII3420CRASH OPIION.............MEIGHT ON OBJECIIVE.... O.DJ IOLERANCE...............PJVOT TOLERANCE.........
NONLINEAR PROBLEMS.

| NONL JNEAR | CONSIRAINTS | 4 | SUPEREASICS LIM |
| :---: | :---: | :---: | :---: |
| NONLINEAR | JACOUIAN VARS | 88 | MESSIAN DIMEMSION. |
| MONL I NEAR | OBJECIIV VARS | 132 | LINESEARCM TOLERANCE |
| PROBLEM N | UMBER. | 0 | REDUCED-GRADIENT POL |

MISCELGANEOUS.
LU ROW TOLERANCE........ 1.00E=03
LU COL TOLERANCE........ 0.10000
LU MOD TOLERANCE A.A.A.E O. 90009

MAJOQ IIERAIIONS LIMIT. MINOR IIERAIIONS LINII』 COMPLETION..............

MUMBER OF MORDS OF CORE AYAILABLE FOR WOAKSPACE 999999

| $\begin{array}{r} 0 \\ 10 \\ 0 \end{array}$ | LOWER BOUND DEFAULIEEP UPPER BDUND DEFAULP:... <br> AIJ TOLERANCE............ | $\begin{aligned} & 0 \cdot 00 E=00 \\ & 3.00 E-00 \\ & 1,00 E-10 \end{aligned}$ |
| :---: | :---: | :---: |
| 10 | (CARD READER)........... | 5 |
| 11 |  | - |
| - | (SCRATCH FILEJ....e.g.0 | - |
| 0 | DUMP FILE.............. | $\bigcirc$ |
| 30 | CYCLE LIMIP.........e.... | 1 |
| 60 | CYCLE TOLERANCEAEsigese |  |
| 1.00E-05 | PARTIAL PRICE FACTORge |  |
| 2,70E-06 | MULTIPLE PRICE.....e.e. |  |
| 2.70E-06 |  |  |
| 135 | DERIVATIVE LEVEL....... | 3 |
| 135 | VERIFY LEVEL.......egoeg | - |
| 0.01000 | OIFFERENCE INTERVAL | 5.39E-06 |
| 0.20000 | CONJUGATE-GRADNT MEIHOD | , |
| 100 | MADIUS OF CONVERGENCE.. | 1.00E-02 |
| 40 |  | 1.00E005 |
| FULL |  |  |
| 0 |  PRINT SPIXES. | $\begin{aligned} & \text { YES } \\ & \text { MO } \end{aligned}$ |
| 1.00E:20 | UNBOUNDED STEP BILELE: | 1.00E+10 |

```
MPS FILE
        I NAME
        2.. ROMS
    101 COLUNNS
gXEX maRNING - ND LINEAR OBJECTIVE FUNCTION FOUND
        5 3 0 ~ R H S ~
        583 BOUNDS
        632 ENDATA
```

MAMES SELECTED


Пнs
RANGES
BOUNDS

RHSI
BOUND:

```044
0
48
``` \(\qquad\)
```

mapalx stailistics

|  | total | NORMAL | FREE | FIXED |
| :---: | :---: | :---: | :---: | :---: |
| nows | 178 | 0 | 0 | 178 |
| COLUMNS | 182 | 2 | 4 | - |

NO. DF MATPIX ELEMENTS 603 DENSITY 1.051

```
NO. DF REJECTED COEFFS 17 A1JTOL ..... 1.00000E-10
```DIGGEST AND SMALLEST COEFFS 0.45090E+00 7.24000E-02lengithof ron-Name hash tagle1409
```

COLLISIONS DURING TABLE LOOKUP ..... 69
NO. OF JACOBIAN ENTRIES SPECIFIED ..... 08
NO. OF LAGHANGE MIULTIPLIERS SPECIFIED ..... 0
ND. OF_ INITIAL _BOUNDS PROCESSED ..... 0
NO. OF SUPERBASICS SPECIFIEO ..... 0
PARTITION SIZE FOR PARTIAC PRICING ..... 182

# BOUNDED <br> 0 <br> 136 

XCLUOING OBJ AND RMSJ



ITN O FEASIBLE SOLUTION. OBJECTIVE E 3.93菏34240E-02

## CHOLESKY FACTOR OF HESSIAN REBET TO I.


START OF MAJOR ITN 3 - PENALIY PARAMETER 2.27E400
ROM ERROR AFTEA RELINEARIIATION E $1.4700 E-11$
RELATIVE CHANGE IN MULIIPLIERS

| Factorize | 3 | demand | 0 | ITEAATION | 3 | INFEAS | , | OBJECTV | 3.49 | 359034E-0 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SLACKS | 19 | LINEAR | 49 | MONLINEAR | 110 | ELEMS | 534 | densily | 1.1 |  |  |  |
| Pa BUMPS | 25 | BPIKE8 | 26 | CORE REOD | 17127 | L LIMIT | 108155 | U LIMIT | 81631 |  |  |  |
| LU BUMPS | 25 | SPIKES | 29 | AIJ ELEMS | 370 | L ELEMS | 302 | $u$ ELEMS | 40 | $F$ ELEMS | 30 | 6.9 |
| Triswaps | 3 | SPKSWP | 0 | REJECTED | 0 | HiN PIV | hatlo | 0.00375 | TOL 3 | 0.00100 | 0.10000 |  |
| ITM 3 | FE | BLE SOLU | N. | BJECIJVE | 3.49 | 90034E-0 |  |  |  |  |  |  |
| NORH RC IS | Alp | SHALL |  | E-09 --- RET | TURN | phase 3. | NOHM | 1. | 053E+0 |  |  |  |
| BIGEEST D | - |  | ORM | -3.952E- | - 09 | PI | 1.053E |  |  |  |  |  |
| END OF maj | R 1 | 3. | TIM | SOLM At MIM | HOR ITM | $0-1$ | OTAL IT | 3 - 3 |  |  |  |  |

BTART OF MAJOR ITN A PENALTY PARAMETER $2.27 E+00$
ROW ERROR AFTER RELINEARIZATION $1.0700 E=11$
RELATIVE CHANGE IN HULTIPLIERS $5.3905 E=09$


objective value 3.4943590339900E-02

LINEAR OBJECTIVE ...
NONLINEAR OBJECTIVE 3.4943590339900E02
PENALTY PARAMETER 0.000000
NORM OF PI $\ldots \ldots \ldots 1.053 \mathrm{I}+01$

NORM OF REDUCED GRADIENT
3.929E-09

ACTIVITY . LOWER LIMIT. . .UPPER LIMIT. .OUAL ACIIVITY ..I
0.00000
$n .00000$
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000

$$
\begin{aligned}
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000 \\
& 0.00000
\end{aligned}
$$







 $00000-0000000000000-0-0000000000000000000000$
 i i





|  | 340 | ROw2P029 | E 0 | 0.00000 | 0.00000 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| A | 105 | MOWZPO30 | $E 0$ | 0.00000 | 0.00000 |
| $\wedge$ | 346 | ROW2P031 | $E 0$ | 0.00000 | 0.00000 |
| A | 347 | ROw2P032 | ED | 0.00000 | 0.00000 |
| - | 348 | ROW2PO33 | EO | 0.00000 | 0.00000 |
| - | 309 | R0w7Pn34 | Ed | 0.00000 | 0.00000 |
| 4 | 350 | ROW2PO35 | EO | 0.00000 | 0.00000 |
| A | 351 | HUW2PO30 | Eu | 0.00000 | 0.00000 |
| A | 352 | ROW2HO37 | E0 | 0.00000 | 0.00000 |
| A | 353 | ROWIPO30 | EO | 0.00000 | 0.00000 |
| A | 354 | ROW2P039 | E0 | 0.00000 | 0.00000 |
| - | 355 | RON2PO40 | 80 | 0.00000 | 0.00000 |
| A | 350 | ROW2PO4I | E0 | 0.00000 | 0.00000 |
| A | 357 | ROW2PO42 | E0 | 0.00000 | 0.00000 |
| A | 350 | ROW2PO43 | E0 | 0.00000 | 0.00000 |
| A | 359 | ROW 2P044 | EO | 0.00000 | 0.00000 |
|  | 500 | ROwnS | C0 | 0.00000 | 0.00000 |
| A | 301 | ROMPS | EO | 0.00009 | 0.00000 |

SECIION 2 - COLUMNS
NUMEER .COLUMN. AT ...ACTIVITY... .OBJ GRADIENT.


| 0.00000 | 0.00000 | -0.01506 | 161 |
| :---: | :---: | :---: | :---: |
| 0.00000 | 0.00000 | 0.00000 | 162 |
| 0.00000 | 0.00000 | 0.00000 | 163 |
| 0.00000 | 0.00000 | 0.00000 | 164 |
| 0.00000 | 0.00000 | 0.00000 | 165 |
| 0.00000 | 0.00000 | 0.00000 | 166 |
| 0.00000 | 0.00000 | 0.00000 | 167 |
| 0.00000 | 0.00000 | 0.00000 | 160 |
| 0.00000 | 0.00000 | 0.00000 | 149 |
| 0.00000 | 0.00000 | 0.00000 | 170 |
| 0.00000 | 0.00000 | -. 00000 | 171 |
| 0.00000 | 0.00000 | 0.00000 | 172 |
| 0.00000 | 0.00000 | 0.00000 | 173 |
| 0.00000 | 0.00000 | 0.00000 | 174 |
| 0.00000 | 0.00000 | 0.00000 | 175 |
| 0.00000 | 0.00000 | 0.00000 | 176 |
| 0.00000 | 0.00000 | -0.25280 | 171 |
| 0.00000 | 0.00000 | 0.00000 | 178 |

.LOWER LIMIT.
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000 0.00000 0.10000 0.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000
3.00000

| 0.00000 | 179 |
| ---: | ---: |
| 0.00000 | 180 |
| 0.00000 | 181 |
| 0.01238 | 182 |
| 0.00000 | 183 |
| 0.00000 | 184 |
| 0.00707 | 185 |
| 0.00000 | 186 |
| .00000 | 187 |
| 0.00000 | 188 |
| 0.00000 | 189 |
| 0.00000 | 190 |
| 0.00000 | 191 |
| 0.00009 | 192 |
| 0.00000 | 193 |
| 0.00712 | 194 |
| 0.00000 | 195 |
| 0.01757 | 196 |
| 0.00000 | 191 |
| 0.04299 | 198 |
| 0.02437 | 199 |
| 0.00000 | 200 |



| 0.00000 | 3.00000 | 0.00000 | 201 |
| :---: | :---: | :---: | :---: |
| 0.00000 | 3.00000 | 0.00000 | 202 |
| 0.00000 | 3.00000 | -8.00000 | 203 |
| 0.00000 | 3.00000 | -0.00000 | 204 |
| 0.00000 | 3.00000 | 0.04917 | 205 |
| 0.00000 | 3.00000 | 0.00000 | 200 |
| 0.00000 | 3.00000 | 0.00000 | 207 |
| 0.00000 | 3.00000 | -.01569 | 208 |
| 0.00000 | 3.00000 | 0.00000 | 200 |
| 0.00000 | 3.00000 | 0.00000 | 210 |
| 0.00000 | 3.00000 | 0.00399 | 211 |
| 0.00000 | 3.00000 | 0.05351 | 212 |
| 0.00000 | 3.00000 | 0.04047 | 213 |
| 0.00000 | 3.00000 | 0.00000 | 214 |
| 0.00000 | 3.00000 | 0.00000 | 215 |
| 0.00000 | 3.00000 | 0.00815 | 210 |
| 0.00000 | 3.00000 | -0.00000 | 217 |
| 0.00000 | 3.00000 | 0.02007 | 216 |
| 0.00000 | 3.00000 | 0.00347 | 219 |
| 0.00000 | 3.00000 | 0.00000 | 220 |
| 0.00000 | 3.00000 | 0.05357 | 221 |
| 0.00000 | 3.00000 | 0.02937 | 222 |
| 0.00000 | 3.00000 | 0.00000 | 223 |
| 0.00000 | 3.00000 | -0.00000 | 224 |
| 0.00000 | 3.00000 | 0.00000 | 225 |
| 0.00000 | 3.00000 | 0.00150 | 226 |
| 0.00000 | 3.00000 | 0.00000 | 221 |
| 0.00000 | 3.00000 | 0.06905 | 220 |
| 0.00000 | 3.00000 | 0.00000 | 229 |
| 0.00000 | 3.00000 | 0.00000 | 230 |
| 0.00000 | 3.00000 | 0.00000 | 231 |
| 0.00000 | 3.00000 | 0.00000 | 232 |
| 0.00000 | 3.00000 | -. 00000 | 233 |
| 0.00000 | 3.00000 | 0.00000 | 234 |
| 0.00000 | 3.00000 | 0.00000 | 235 |
| 0.00000 | 3.00000 | 0.00000 | 236 |
| 0.00000 | 3.00000 | 0.00000 | 237 |
| 0.00000 | 3.00000 | 0.00000 | 238 |
| 0.00000 | 3.00000 | -. 00000 | 239 |
| 0.00000 | 3.00000 | -. 00000 | 240 |
| 0.00000 | 3.00000 | 0.00000 | 241 |
| 0.00000 | 3.00000 | 0.00000 | 242 |
| 0.00000 | 1.00000 | 0.00000 | 243 |
| 0.00000 | 1.00000 | 0.00000 | 244 |
| 0.00000 | 3.00000 | -0.00000 | 245 |



$\square$







 $0000^{\circ}$




| 161 | COLmuO29 | 85 | 0.94256 | 0.00000 |
| :---: | :---: | :---: | :---: | :---: |
| 162 | colmuoso | B 3 | 0.98991 | 0.00000 |
| 103 | COLMUO31 | H | 0.98991 | 0.00000 |
| 104 | COLMUO32 | BS | 0.98993 | 0.00000 |
| 165 | COLMUO33 | 83 | 0.98991 | 0.00000 |
| 166 | COLMUO3A | 83 | 0.98993 | 0.00000 |
| 167 | colmuoss | HS | 0.90993 | n.00000 |
| 168 | colmuosb | BS | 0.98993 | 0.00000 |
| 169 | COLMUO37 | - 5 | 0.85220 | 0.00000 |
| 170 | colmuose | 83 | 0.93935 | 0.00000 |
| 171 | colmuos | BS | 0.98993 | 0.00000 |
| 172 | COLmunas | B | 0.98993 | 0.00000 |
| 173 | CULMUNAI | B 5 | 0.98993 | 0.00000 |
| 170 | colmuoaz | 85 | 0.98991 | 0.00000 |
| 175 | Colmuas 3 | us | 0.98993 | 0.00000 |
| 170 | COLmU044 | -S | 0.98993 | 0.00000 |
| 171 | CL2K1 | Sus | 0.70338 | 0.00000 |
| 170 | CL2K2 | OS | 0.17950 | 0,00000 |
| 179 | CL2x3 | HS | -0.17950 | 0.00000 |
| 180 | CL2H1 | HS | 0.64425 | 0.00000 |
| 181 | CL2P2 | 85 | 0,09361 | 0.00000 |
| 182 | CL2P3 | $\bullet 5$ | -0.09361 | 0.00000 |
| 183 | RHSI | EO | -1.00000 | 0.00000 |



## $\stackrel{\sim}{\infty}$

PROQLEM NAME
STATUS

OBJECTIVE
RHS
MANGES
©OUNOS

## (MIN)

RHS!
$\triangle$ OUNDI

SECTION 1 - ROWS

|  | 184 | NOWSOO8 | E0 | 0.00000 | 0.00000 |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 185 | ROw300ट | E0 | 0.00000 | 0.00000 |
|  | 180 | R0w3003 | EO | 0.00000 | 0.00000 |
|  | 181 | HUnS004 | EO | 0.00000 | 0.00000 |
| d | 184 | ROwS 005 | B | 0.00000 | 0.00000 |
|  | 189 | Rums006 | EO | 0.00000 | 0.00000 |
|  | 190 | RUwS007 | EO | 0.00000 | 0.00000 |
|  | 191 | ROwS008 | E0 | 0.00000 | 0.00000 |
|  | 192 | KUwSOU9 | EO | 0.00000 | 0.00000 |
| 0 | 193 | HOwSOIU | HS | 0.00000 | 0.00000 |
|  | 196 | RONSU11 | E0 | 0.00000 | 0.00000 |
|  | 195 | HOwSO12 | E0 | 0.00000 | 0.00000 |
| D | 190 | HUwS015 | 83 | 0.00000 | 0.00000 |
|  | 197 | HOwSO19 | B | 0.00000 | 0.00000 |
|  | 190 | ROw 3015 | E0 | 0.00000 | 0.00000 |
|  | 199 | ROW3016 | EO | 0.00000 | 0.00000 |
|  | 200 | HUw3017 | E0 | 0.00000 | 0.00000 |
| D | 201 | HUWS018 | B 5 | 0.00000 | 0.00000 |
|  | 202 | ROws 019 | E0 | 0.00000 | 0.00000 |
| 0 | 203 | HOW 9020 | B | 0.00000 | 0.00000 |
|  | 204 | H0wS021 | E0 | 0.00000 | 0.00000 |
|  | 205 | HUmS022 | E0 | 0.00000 | 0.00000 |
| D | 206 | H0NS023 | BS | 0.00000 | 0.00000 |
|  | 201 | ROM 8024 | EO | 0.00000 | 0.00000 |
|  | 200 | H0w 3025 | E 0 | 0.00000 | 0.00000 |
|  | 209 | ROWS026 | ED | 0.00000 | 0.00000 |
|  | 210 | RUnSU27 | E ${ }^{\text {d }}$ | 0.00000 | 0.00000 |
|  | 211 | MOmS028 | EO | 0.00000 | 0.00000 |
| 0 | 212 | HOwS 029 | BS | 0.00000 | 0.00000 |
|  | 213 | HOmS030 | E 4 | 0.00000 | 0.00000 |
|  | 314 | H0wS031 | EO | 0.00000 | 0.00000 |

## non-substitution problem

$2.6793413649 E=02$
superbasics
-•LOWER LIMIT...UPPER LIMIT. .DUAL ACTIVITY ••.J

| 0.00000 | 0.00000 | -0.01420 | 1 |
| :---: | :---: | :---: | :---: |
| 0.00000 | 0.00000 | 0.15127 | 1 |
| 0.00000 | 0.00000 | -0.60761 | 1 |
| 0.00000 | 0.00000 | -0.14004 | , |
| 0.00000 | 0.00000 | 0.00000 | 5 |
| 0.00000 | 0.00000 | -0.06905 | + |
| 0.00000 | 0.00000 | -0.56775 | 1 |
| 0.00000 | 0.00000 | 0.15265 |  |
| 0.00000 | 0.00000 | -0.08150 | - |
| 0.00000 | 0.00000 | 0.00000 | 10 |
| 0.00000 | 0.00000 | -0.33926 | 11 |
| 0.00000 | 0.00000 | -0.41574 | 12 |
| 0.00000 | 0.00000 | 0.00000 | 13 |
| 0.00000 | 0.00000 | 0.00000 | 1 |
| 0.00000 | 0.00000 | 0.01818 | 15 |
| 0.00000 | 0.00000 | -0.02284 | 16 |
| 0.00000 | 0.00000 | -0.17498 | 11 |
| 0.00000 | 0.00000 | 0.00000 | 11 |
| 0.00000 | 0.00000 | -0.62063 | 19 |
| $0.00000^{\circ}$ | 0.00000 | 0.00000 | 20 |
| 0.00000 | 0.00000 | -0.54982 | 21 |
| 0.00000 | 0.00000 | -0.49671 | 22 |
| 0.00000 | 0.00000 | 0.00000 | 13 |
| 0.00000 | 0.00000 | -0.43521 | 24 |
| 0.00000 | 0.00000 | -0.21652 | 25 |
| 0,00000 | 0.00000 | 0.00004 | 26 |
| 0.00000 | 0.00000 | -0.09373 | 27 |
| 0.00000 | 0.00000 | -0.38230 | 20 |
| 0.00000 | 0.00000 | 0.00000 | 29 |
| 0.00000 | 0.00000 | -0.10506 | 30 |
| 0.00000 | 0.00000 | 0.00064 | 31 |



























PROBLEM NAME
STATUS
OBJECTIVE
OPTIMAL SOLN

## (MIN)

RHS
RANGES
BOUNDS

OBJECTIVE VALUE
ITEHATION 12

RHS
BOUNDI
SECTION 1- MOW8


# non-substitution problem 

3.5142842356E=02

SUPEHBASICS 4

| 0.00000 | 0.00000 | -2.31841 | 1 |
| :---: | :---: | :---: | :---: |
| 0.00000 | 0.00000 | -0.00970 | 2 |
| 0.00000 | 0.00000 | -0.04101 | 3 |
| 0.00000 | 0.00000 | 0.04296 | 4 |
| 0.00000 | 0.00000 | -0.23524 | 5 |
| 0.00000 | 0.00000 | -0.01241 | 6 |
| 0.00000 | 0.00000 | 0.00000 | 7 |
| 0.00000 | 0.00000 | 0.12728 | $\sigma$ |
| 0.00000 | 0.00000 | 0.00000 | 9 |
| 0.00000 | 0.00000 | -0.11795 | 10 |
| 0.00000 | 0.00000 | -0.09589 | II |
| 0.00000 | 0.00000 | 0.03461 | 12 |
| 0.00000 | 0.00000 | -0.00889 | 13 |
| 0.00000 | 0.00000 | 0.07279 | 15 |
| 0.00000 | 0.00000 | 0.06114 | 15 |
| 0.00000 | 0.00000 | 0.00000 | 16 |
| 0.00000 | 0.00000 | - 0.12886 | 17 |
| 0.00000 | 0.00000 | 0.04554 | 18 |
| 0.00000 | 0.00000 | 0.00000 | 19 |
| 0.00000 | 0.00000 | 0.02670 | $20^{-}$ |
| 0.00000 | 0.00000 | -0.20797 | 21 |
| 0.00000 | 0.00000 | -0.15608 | 22 |
| 0.00000 | 0.00000 | 0.64056 | 23 |
| 0.00000 | 0.00000 | -0.26133 | 24 |
| 0.00000 | 0.00000 | -0.03167 | 25 |
| 0.00000 | 0.00000 | -0.24052 | $26^{-}$ |
| 0.00000 | 0.00000 | 0.02099 | 27 |
| 0.00000 | 0.00000 | 0.00000 | 28 |
| 0.00000 | 0.00000 | 0.12470 | 29 |
| 0.00000 | 0.00000 | 0.03769 | 30 |
| 0.00000 | 0.00000 | -0.12623 | 31 |






mmmmmmmmmmmmmmmmmmmnmmmmmmmmmmmmmmmmammmmamm


ḿcommomacmmmmomm



|  | 261 | ROmYos4 | E ${ }^{\text {a }}$ | 0.96470 |
| :---: | :---: | :---: | :---: | :---: |
|  | 262 | HOnYO35 | EG | 0.86240 |
|  | 263 | 日0wroso | E0 | 0.90090 |
|  | 264 | HOwYOS1 | E0 | 0.65220 |
|  | 265 | ROWYOS8 | EO | 0.85560 |
|  | 266 | RUwYose | Eu | 0.98040 |
|  | 267 | RUwY040 | ED | 0.41080 |
|  | 268 | HOwYOA1 | EO | 0.90520 |
|  | 269 | ROwrome | E0 | 0.94490 |
|  | 270 | NOMros3 | EU | 0.92940 |
|  | 211 | MOWYOA4 | EO | 0.95230 |
|  | 272 | ROm2K001 | EO | 0.00000 |
| 1 | 275 | ROmZKOU? | EO | 0.00000 |
| A | 274 | ROW2K0U3 | EU | 0.00000 |
| A | 275 | RUW2K004 | EO | 0.00000 |
| 1 | 276 | ROW2K005 | E0 | 0.00000 |
| a | 271 | ROW2K006 | E0 | 0.00000 |
| 4 | 278 | RUm2nu0) | E 4 | 0.00000 |
| , | 279 | ROWZKOUS | E0 | 0.00000 |
| A | 280 | ROWZK009 | EO | 0.00000 |
| A | 281 | ROW2x010 | E 4 | 0.00000 |
| 1 | 282 | ROWZK011 | E0 | 0.00000 |
| 1 | 205 | HONZKO12 | $E 0$ | 0.00000 |
| A | 204 | HOw2K013 | E0 | 0.00000 |
| A | 285 | ROw2K014 | Eu | 0.00000 |
|  | 286 | HOw2K015 | EO | 0.00000 |
|  | 201 | ROW2K016 | ED | 0.00000 |
| 1 | 288 | HU*2K017 | E0 | 0.00000 |
| 1 | 209 | HUw2K016 | E0 | 0.00000 |
| A | 290 | ROW2K019 | E0 | 0.00000 |
| A | 291 | ROW2M020 | EO | 0.00000 |
| A | 292 | ROW2K021 | EO | 0.00000 |
| A | 293 | ROW2K022 | E 0 | 0.00000 |
|  | 294 | ROWZKO23 | EO | 0.00000 |
| A | 295 | ROW2K024 | E0 | 0.00000 |
| A | 296 | ROWIK025 | E0 | 0.00000 |
| A | 291 | ROW2K026 | EO | 0.00000 |
| - | 290 | ROW2K027 | EO | 0.00000 |
| A | 299 | ROW2K020 | E0 | 0.00000 |
| i | 300 | ROW2K029 | E0 | 0.00000 |
| 1 | 301 | ROW2K030 | E 0 | 0.00000 |
| , | 302 | ROW2KO34 | $E 0^{-}$ | 0.00000 |
| 1 | 303 | RDW2K032 | EO | 0.00000 |
| 4 | 304 | ROW2K033 | $E 0$ | 0.00000 |









|  | 350 | ROM2POS5 | EU | 0.00000 | 0.00000 |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 351 | HOWZPOSO | EO | 0.00000 | 0.00000 |
| A | 352 | HUm2P037 | EU | 0.00000 | 0.00000 |
| 1 | 355 | HOW2PO36 | E0 | 0.00000 | 0.00000 |
|  | 354 | H0w2PO39 | E0 | 0.00000 | 0.00000 |
|  | 355 | RUW2H040 | E0 | 0.00000 | 0.00000 |
|  | 350 | HON2PO41 | EU | 0.00000 | 0.00000 |
|  | 351 | ROW2P042 | EO | 0.00000 | 0.00000 |
|  | 356 | ROW2P043 | E0 | 0.00000 | 0.00000 |
|  | 359 | ROW2P044 | E0 | 0.00000 | 0.00000 |
| $\wedge$ | 360 | RUWKS | Ev | 0.00000 | 0.00000 |
|  | 361 | ROmPS | E0 | 0.00000 | 0.00000 |

## SECTION 2 - COLUMNS

| NUMBER |  | .COLUMN. | 4 | ...ACTIVITY... | . OB J | GRADIENT. |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| D | 1 | COLSKOOI | 85 | 0.00000 |  | 0.00000 |
|  | 2 | COLSK002 | 85 | 0.36929 |  | 0.00000 |
|  | 3 | COLSKOUS | 83 | 0.35276 |  | 0.00000 |
|  | 4 | COLSK004 | 88 | 0.22156 |  | -0.00000 |
|  | 5 | COLSK005 | BS | 0.22759 |  | -0.00000 |
|  | 6 | COLSK006 | 85 | 0.36384 |  | -0.00000 |
|  | 1 | COLSK007 | 88 | 0.28302 |  | 0.00000 |
|  | * | COLSK008 | 85 | 0.22754 |  | 0.00000 |
|  | 9 | COLSK0U9 | 85 | 0.22756 |  | 0.00000 |
|  | 10 | COLSKO10 | 63 | 0.22754 |  | 0.00000 |
|  | 11 | COLSKOLI | 85 | 0.22760 |  | 0.00000 |
|  | 12 | COLSK012 | 883 | 0.22756 |  | 0.00000 |
|  | 13 | COLSMOLS | 85 | 0.22760 |  | 0.00000 |
|  | 14 | COLSK014 | 88 | 0.22754 |  | 0.00000 |
| 0 | 15 | COLSXOI5 | 85 | 0.00000 |  | 0.00000 |
|  | 16 | COLSKOI6 | LL | 0.00000 |  | 0.00000 |
|  | 17 | COLSKOI7 | BS | 0.22756 |  | 0.00000 |
|  | 18 | COLSK018 | 85 | 0.22154 |  | 0.00000 |
|  | 10 | COLSKO19 | 88 | 0.22160 |  | 0.00000 |
|  | 20 | COLSK020 | B8 | 0.22156 |  | 0.00000 |
|  | 21 | COLSK021 | B8 | 0.22756 |  | 0.00000 |
|  | 22 | COLSK022 | 85 | 0.22760 |  | 0.00000 |
| D | 23 | COLSK023 | 83 | 0.00000 |  | 0.00000 |
|  | 24 | COLSK024 | BS | 0.29015 |  | 0.00000 |
|  | 25 | COLSK025 | 88 | 0.26164 |  | 0.00000 |
|  | 26 | COL8K026 | B | 0.22759 |  | 0.00000 |
|  | 21 | COL8K027 | 88 | 0.22760 |  | 0.00000 |
|  | 23 | COLSM020 | B8 | 0.32273 |  | 0.00000 |

0,00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000
0.00000

| 0.00000 | -0.06032 | 167 |
| :--- | ---: | :--- |
| 0.00000 | -0.02182 | 168 |
| 0.00000 | 0.00000 | 164 |
| 0.00000 | 0.00000 | 170 |
| 0.00000 | 0.06368 | 171 |
| 0.00000 | -0.01192 | 172 |
| 0.00000 | -0.01752 | 173 |
| 0.00000 | 0.02218 | 174 |
| 0.00000 | 0.00668 | 175 |
| 0.00000 | 0.02956 | 176 |
| 0.00000 | -0.00000 | 177 |
| 0.00000 | -0.25171 | 178 |

OWER LIMIT. ..UPPER LIMIT. REUUCED GRADNT M\&J

| 0.00000 | 1.00000 | 0.00000 | 179 |
| :---: | :---: | :---: | :---: |
| 0.00000 | 1.00000 | $=0.00000$ | 180 |
| 0.00000 | 1.00000 | -0.00000 | 181 |
| 0.00000 | 1.00000 | -0.00000 | 182 |
| 0.00000 | 1.00000 | $=0.00000$ | 183 |
| 0.00000 | 1.00000 | -0.00000 | 184 |
| 0.00000 | 1.00000 | 0.00000 | 185 |
| 0.00000 | 1.00000 | 0.00000 | 186 |
| 0.00000 | 1.00000 | 0.00000 | 101 |
| 0.00000 | 1.00000 | 0.00000 | 188 |
| 0.00000 | 1.00000 | 0.00000 | 189 |
| 0.00000 | 1.00000 | 0.00000 | 190 |
| 0.00000 | 1.00000 | 0.00000 | 191 |
| 0.00000 | 1.00000 | 0.00000 | 192 |
| 0.00000 | 1.00000 | 0.00000 | 193 |
| 0.00000 | 1.00000 | 0.10169 | 194 |
| 0.00000 | 1.00000 | 0.00000 | 195 |
| 0.00000 | 1.00000 | 0.00000 | 196 |
| 0.00000 | 1.00000 | 0.00000 | 197 |
| 0.00000 | 1.00000 | 0.00000 | 198 |
| 0.00000 | 1.00000 | 0.00000 | 199 |
| 0.00000 | 1.00000 | 0.00000 | 200 |
| 0.00000 | 1.00000 | 0.00000 | 201 |
| 0.00000 | 1.00000 | 0.00000 | 202 |
| 0.00000 | 1.00000 | 0.00000 | 203 |
| 0.00000 | 1.00000 | 0.00000 | 204 |
| 0.00000 | 1.00000 | 0.00000 | 205 |
| 0.00000 | 1.00000 | 0.00000 | 206 |




| 0 | 74 | COLSP030 | B 3 | 0.00000 | 0.00000 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 0 | 75 | COLSP031 | B S | 0.00000 | 0.00000 |
| 0 | 76 | COLSPO32 | 88 | 0.00000 | 0.00000 |
| 0 | 11 | COLSPOS3 | 83 | 0.00000 | 0.00000 |
|  | 76 | COLSPO34 | L6 | 0.00000 | 0.00000 |
| 0 | 19 | CULSP035 | 48 | 0.00000 | 0.00000 |
| U | 60 | COLSPOS0 | B 8 | 0.00000 | 0.00000 |
|  | 01 | COLSP037 | BS | 0.15610 | 0.00000 |
|  | 82 | COLSr038 | BS | 0.05230 | U.00000 |
|  | 83 | COLSPO34 | LL | 0.00000 | 0.00000 |
| U | 84 | COLSPO40 | B 8 | 0.00000 | 0.00000 |
| U | 85 | colspual | B S | 0.00000 | U.00000 |
| 0 | 86 | COLSPOM2 | B8 | 0.00000 | 0.00000 |
| 0 | 61 | COLSPO43 | 83 | 0.00000 | 0.00000 |
| 0 | 08 | COLSPO44 | 83 | 0.00000 | 0.00000 |
|  | 09 | COLEEOOI | B8 | -0.03454 | -0.03454 |
|  | 90 | COLEEOU2 | 88 | -0.0035 | -0.00358 |
|  | 91 | COLEEOUS | H3 | -0.01447 | -0.01447 |
|  | 92 | COLEEOUA | BS | 0.00978 | 0.00978 |
|  | 93 | COLEEOOS | BS | -0.05354 | -0.05350 |
|  | 94 | COLEEOOO | 83 | -0.00460 | -0.00468 |
|  | 45 | COLEEOOT | - 3 | 0.05344 | 0.05544 |
|  | 96 | COLEEOUA | B | 0.02896 | 0.02896 |
|  | 97 | COLEEOOQ | HS | 0.03088 | 0.03888 |
|  | 98 | COLEEOIO | 48 | -0.02684 | -0.02084 |
|  | 99 | CULEE011 | H8 | -0.02182 | -0.02182 |
|  | 100 | COLEEOI2 | 88 | 0.00188 | 0.00788 |
|  | 101 | COLEEOIS | 4S | -0.00202 | -0.00202 |
|  | 102 | CULEEO14 | 89 | 0.01656 | 0.01056 |
|  | 103 | CULEE 015 | HS | 0.01403 | 0.01463 |
|  | 104 | COLEEUS 6 | H8 | 0.10169 | 0.10169 |
|  | 105 | COLEEOI7 | 88 | -0.02932 | -0.02932 |
|  | 106 | COLEEO ${ }^{\circ}$ | B8 | 0.01036 | 0.01036 |
|  | 101 | COLEEUI9 | US | 0.03088 | 0.03888 |
|  | 108 | COLEEOCO | BS | 0.00608 | 0.00608 |
|  | 104 | COLEE021 | B 3 | -0.04132 | -0.04732 |
|  | 110 | CULEE022 | H8 | -0.03552 | -0.03552 |
|  | 111 | COLEEUS3 | BS | 0.04726 | 0.04726 |
|  | 112 | COLEEUZ4 | HS | -0.07687 | -0.07687 |
|  | 113 | COLEE025 | 83 | -0.00026 | -0.00620 |
|  | 114 | COLEEOL6 | HS | -0.05474 | -0.05474 |
|  | 115 | CULEE 021 | as | 0.00476 | 0.00478 |




| none | NONE | 0.00000 | 290 |
| :---: | :---: | :---: | :---: |
| NONF | NONE | 0.00000 | 295 |
| NONE | NONE | 0.00000 | 296 |
| NONE | NONE | 0.00000 | 297 |
| NUNE | NONE | 0.00000 | 298 |
| NONE | NONE | 0.00000 | 299 |
| NUNE | NONE | 0.00000 | 300 |
| NONE | none | 0.00000 | 301 |
| NONE | NONE | 0.00000 | 302 |
| NONE | NONE | 0.00000 | 303 |
| NUNE | NUNE | 0.00000 | 304 |
| NUNE | NONE | 0.00000 | 305 |
| NONE | NONE | 0.00000 | 306 |
| none | NDNE | 0.00000 | 307 |
| NONE | nune | 0.00000 | 308 |
| NONE | NONE | 0.00000 | 309 |
| NONE | NONE | 0.00000 | 310 |
| 0.00000 | 1. 00000 | 0.00000 | 311 |
| 0.00000 | 1.00000 | 0.00000 | 312 |
| 0.00000 | 1.00000 | 0.00000 | 313 |
| 0.00000 | 1.00000 | 0.00000 | 314 |
| 0.00000 | 1.00000 | 0.00000 | 315 |
| 0.00000 | 1.00000 | 0.00000 | 316 |
| 0.00000 | 1.00000 | 0.00000 | 317 |
| 0.00000 | 1.00000 | -0.00000 | 315 |
| 0.00000 | . 00000 | 0.00000 | 319 |
| 0.00000 | 1.00000 | $-0.00000$ | 320 |
| 0.00000 | . 00000 | 0.00000 | 321 |
| 0.00000 | . .00000 | 0.00000 | 322 |
| 0.00000 | . 00000 | 0.00000 | 323 |
| 0.00000 | 1.00000 | 0.00000 | 324 |
| 0.00000 | . .00000 | 0.00000 | 325 |
| 0.00000 | . 00000 | 0.00000 | 326 |
| 0.00000 | . 00000 | 0.00000 | 321 |
| 0.00000 | . 00000 | 0.00000 | 320 |
| 0.00000 | . 00000 | 0.00000 | 329 |
| 0.00000 | . 00000 | 0.00000 | 330 |
| 0.00000 | . 00000 | 0.00000 | 331 |
| 0.00000 | . 00000 | 0.00000 | 332 |
| 0.00000 | . 00000 | $-0.00000$ | 333 |
| 0.00000 | . 00000 | 0.00000 | 334 |
| 0.00000 | . 00000 | 0.00000 | 335 |
| 0.00000 | . 00000 | 0.00000 | 336 |
| 0.00000 | . 00000 | 0.00000 | 337 |
| 0.00000 | 1.00000 | 0.00000 | 338 |





##  <br> 

Anderson, R. L. and L. A. Nelson. [1975], "A Family of Models Involving Intersecting Straight Lines and Concomitant Experimental Designs Useful in Evaluating Response to Fertilizer Nutrients," Biometrics 31:303-18.

Arndt, W. and G. A. McIntyre. [1963], "The Initial and Residual Effects of Superphosphate and Rock Phosphate for Sorghum on a Lateritic Red Earth," Australian J. Agric. Research, 14:785-795.

Atkinson, A. C. [1970], "A Method for Discriminating Between Models," Journal of the Royal Statistical Society," 32 Series B.

Balmukand, B. H. [1928], Studies in Crop Variation - V. "The Relation Between Yield and Soil Nutrients," Journal of Agricultural Science 18(4):602-627.

Barber, S. A. [1958], "The Relation of Fertilizer Placement to Nutrient Uptake and Crop Yield: I. The Interaction of Row Phosphorus and Soil Level of Phosphorus," Agronomy Journal 50:535-539.
_ [1979], "Soil Phosphorus after 25 Years of Cropping with Five Rates of Phosphorus Application," Commun. in Soil Sci. and Plant Analysis 10(11):1459-1468.
. [1980], "Twenty-Five Years of Phosphate and Potassium Fertilization of a Corn Crop," Fertilizer Research 1:29-36.

Baule, B. [1918], "Zu Mitscherlichs Gesetz der Physiologischen Beziehungen," Landwirtschaftliche Jahrbucher 51:363-385.

Baum, E. L., E. O. Heady and J. Blackmore (Eds.) [1956], "Methodological Procedures in the Economic Analysis of Fertilizer Use Data," Ames, Iowa State College Press.

Baum, E. L., E. O. Heady, J. T. Pesek and C. G. Hildreth (Eds.) [1957], "Economic and Technical Analysis of Fertilizer Innovation and Resource Use," Ames, Iowa State College Press.

Boyd, D. A. [1970], "Some recent ideas on fertilizer response curves," Proceedings Congress Int. Potash Inst., Berne, on the Role of Fertilization in the Intensification of Agricultural Production Sept. 15-18, Antibes.

Bray, R. H. [1954], "A Nutrient Mobility Concept of Soil-Plant Relationships," Soil Science 78(1):9-23.
. [1963], "Confirmation of the Nutrient Mobility Concept of Soil-Plant Relationships," Soil Science 95:124-129.

Brown, G. W. and M. N. Overson. [1958], "Production Functions from Data

Over a Series of Years," JFE 40:451-457.
Brown, W. G. et al. [1956], "Production Functions, Isoquants, Isoclines and Economic Optima in Corn Fertilization for Experiments with 2 or 3 Variable Nutrients," Iowa State College Ag. Expt. Statn. Bull. 441.

Colwell, J. D. and R. J. Esdaile. [1966], "The Application of Production Function Analysis for the Estimation of Fertilizer Requirements of Wheat in Northern New South Wales," Aust. Jour. Expt. Agric. Animal Husb. 6:418-424.

Colyer, D. and E. M. Kroth. [1968], "Corn Yield Response and Economic Optima for Nitrogen Treatments and Plant Population over a SevenYear Period," Agronomy Journal 60(5):524-529.

Cope, J. T. and R. D. Rouse. [1973], "Interpretation of Soil Test Results, Soil Testing and Plant Analysis," Eds. Walsh and Beaton, Revised Ed. Soil Sci. Society of America Inc., Madison, Wisconsin.

Cox, D. R. [1961], "Tests of Separate Families of Hypotheses" Proceedings of the Fourth Berkeley Symposium I.

Davidson, R. and J. G. MacKinnon. [1981], "Several Tests for Model Specifications in the Presence of Alternative Hypotheses," Econometrics 49(3):781-793.

Dhrymes, P. J. et al. [1970], "Estimation of Distributed Lags," International Economic Reviews 11(2):235-250.

Dhrymes, P. J. [1971], "On the Strong Consistency of Estimators for Certain Distributed Lag Models with Autocorrelated Errors," International Economic Review 12(2):329-343. - [1971], Distributed Lags: Problems of Estimation and Formulation, Holden-Day Inc. San Francisco.

Engelstad, 0. P. [1963], "Effect of Variation in Fertilizer Rates and Ratios on Yield and Profit Surfaces," Agronomy Journal 55:263-265.

Engelstad, O. P. and G. L. Terman. [1966], "Fertilizer Nitrogen: Its Role in Determining Crop Yield Levels," Agronomy Journal 58:536-539.

Engelstad, O. P. and W. L. Parks. [1971], "Variability in Optimum Nitrogen Rates for Corn," Aust. J. Agr. Econ. 13:35-46.
F.A.O. [1952], Fertilizers - A World Report on Production and Consumption, Rome.
$\qquad$ - [1979], Fertilizer Yearbook, Vol. 29, Rome.

Fisher, G. R. and M. McAleer. [1981], "Alternative Procedure and Associated Tests of Significance for Non-nested Hypotheses," Journal
of Economics, May 1981.
Fixen, P. E. and P. L. Carson. [1978], "The Relationship Between Soil Test and Small Grain Response to P-Fertilization in Field Experiments," Agronomy Journal 70:838-844.

French, B. L. [1956], "Functional Relationships for Irrigated Corn Responses to Nitrogen," JFE 38:736-747.

Fuller, W. A. [1965], "Stochastic Fertilizer Production Functions for Continuous Corn," JFE 47:105-119.

Gaver, K. and M. S. Geisel. [1974], "Discriminating Among Alternative Models: Bayesian and Non-Bayesian Methods" in P. Zarembka, Frontiers in Econometrics, Academic Press, New York.

Havlicek, J., Jr. and J. A. Seagraves. [1962], The "Cost of the Wrong Decision" as a Guide in Production Research, JFE 44:157-168.

Helyar, K. R. and D. P. Godden. [1977], "The Biology and Modelling of Fertilizer Response," Jour. Aust. Inst. Agr. Sci. 43:22-30.

Heady, E. O. and J. Pesek. [1954], "A Fertilizer Production Surface with Specification of Economic Optima for Corn Grown Calcareous Ida Silt Loam," JFE 36:466-482.

Heady, E. O. and W. G. Brown. [1955], "Crop Response Surfaces and Economic Optima in Fertilizer Use," Iowa State College Ag. Expt. Statn. Bull. 424.

Heady, E. O. and Rao. [1956], "Fertilizer Production Functions from Experimental Data with Associated Supply and Demand Relationships," Iowa State Univ. Ag. Home Econ. Exp. Sta. Res. Bull. 543, Ames.

Hildreth, C. G. [1954], "Point Estimates of Ordinates of Concave Functions" I. American Stat. Assoc. 49:598-619. - [1956], "Discrete Models with Qualitative Restrictions," in Baum, E. L., Heady, E. O. and Blackmore, J. (Eds.). - [1957], "Possible Models for Agronomic-Economic Research" in Baum, E. L. et al. (Eds.).

Hunway, J. and L. Dumenil. [1955], "Predicting Nitrogen Needs of Iowa Soils: III. Use of Nitrate Production Together with Other Information as a Basis for Making Nitrogen Fertilizer Recommendations for Corn in Iowa," Soil Sci. Soc. Amer. Proc. 19:17-80.

Hutton, R. F. [1955], "An Appraisal of Research on Economics of Fertilizer Use," TVA Report \#T55-1.

Intriligator, M. D. [1978], Economic Models, Techniques and Applications, Prentice-Hall, Inc., Engelwood-Cliffs.

Jensen, D. and J. Pesek. [1959a \& b], "Generalization of Yield Equations in Two or More Variables," I and II, Agronomy Journal 51:255-263.

Johnson, G. L. [1957], "Planning Agronomic-Economic Research in View of Results to Date, and the Role of Management in Planning Farms for Optimum Fertilizer Use," in Baum, E. L., et al. 217-225, 261270.

Kennedy, J. S.et al. [1973], "Optimal Fertilizer Carryover and Crop Recycling Policies for a Tropical Grain Crop," Aust. Jour. Agric. Econ. 17:104-113.

Klein, L. R. [1958], "The Estimation of Distributed Lags," Econometrica 26(4):553-565.

Lanzer, E. A. [1978], "Fertilizer Recommendations from the Dynamic Liebig-Mitscherlich Model: The Case of Wheat-Soybeans in Southern Brazil," Unpub. Ph.D. Dissertation, Ag. Econ. Dept., University of California, Davis.

Lanzer, E. A. and Q. Paris. [1979], "The Dynamic Liebig-Mitscherlich Model for Optimal Fertilizer Recommendations," Unpub. Manuscript, Ag. Econ. Dept., University of California, Davis.

Liebig, J. von. [1840], "Organic Chemistry in its Application to Agriculture and Physiology," London, Playfair.

Liviatan, N. [1965], "Estimates of Distributed Lag Consumption Functions from Cross-Section Data," Review of Economics and Statistics 47:4453.

Maddala, G. S. [1977], Econometrics, McGraw-Hill, Inc., New York.
Matocha, J. E. et al. [1970], "Residual Value of Phosphorus Fertilizer on a Calcareous Soil," Agronomy Journal 62:572-574.

Mitscherlich, E. A. [1909], "Das Gestz des Minimums und das Gesetz des abnehmenden Bodenertrages." Landwirtschaftliche Jahrbucher, 38:537552.

Mombiela, F. A. and L. A. Nelson. [1980], "Evaluation of Fertilizer Recommendation Methodologies Using Simulated Corn P Responses in North Carolina Norfolk-Like Soils with Varying Soil P Fertility," North Carolina Ag. Research Technical Bull. \#28.

Munson, R. D. and J. P. Doll. [1959], "The Economics of Fertilizer Use in Crop Production," Advances in Agronomy 11:133-169.

Nelder, J. A. [1966], "Inverse Polynomials, a Useful Group of Multifactor Response Functions," Biometrics 22:128-141.

Perrin, R. K. [1976], "The Value of Information and the Value of Theoretical Models in Crop Response Research," AJAE 58(1):54-61.

Pesaran, M. H. [1973], "The Small Sample Problem of Truncation Remainders in The Estimation of Distributed Lag Models with Autocorrolated Errors," International Economic Review 14(1):120-131. - [1974], "On the General Problem of Model Selection," Review of Econ. Studies 41(2):153-172.

Pesaran, M. H. and A. S. Deaton. [1978], "Testing Non-nested Non-linear Regression Models," Econometrica 46(3):677-694.

Pesek, J. and E. D. Heady. [1952], "A Two Nutrient-Response Function with Determination of Economic Optima for the Rate and Grade of Fertilizer for Alfalfa," Soil Sci. Society of America: Proceedings 20: 240-246.

Pesek, J. T. [1956], "Agronomic Problems in Securing Fertilizer Response Data Desirable for Economic Analysis," in Baum, E. L., Heady, E. O., and Blackmore, J. (Eds.).

Poirier, D. J. [1973], "Piecewise Regression Using Cubic Splines," Journal of the American Stat. Assoc. 68(343):515-524.

Putter, J. et al. [1966], "Quadratic Equations as an Interpretative Tool in Biological Research," Agron. Jour. 58:103-104.

Quandt, R. E. [1974], "A Comparison of Methods for Testing Non-nested Hypotheses," Review of Economics and Statistics 56:92-99.

Reuss, J. 0. and J. M. Geist. [1970], "Prediction of the Nitrogen Requirements of Nitrogen Release," Agronomy Journal 62:381-389.

Russel, E. J. [1950], Soil Conditions and Plant Growth, 8th Ed. Longmans, New York.

Ryan, J. G. [1972], A Generalized Crop Fertilizer Production Function, Unpub. Ph.D. Dissertation, North Carolina State University at Raleigh.

Smith, P. L. [1979], "Splines as a Useful and Convenient Statistical Tool," The American Statistician 33(2):57-62.

Spillman, W. J. [1923], "Application of the Law of Diminishing Returns to Some Fertilizer and Feed Data," Journal of Farm Economics 5(1): 36-52.

Stauber, M. S. and O. R. Burt. [1973], "Implicit Estimate of Residual Nitrogen Under Fertilized Range Conditions in the Northern Great Plains," Agronomy Journal 65:897-901.

Stauber, M. S., O. R. Burt, and F. Linse. [1975], "Economic Evaluation of Nitrogen Fertilization of Grasses When Carryover is Significant," AJAE 57(3):463-471.

Stemberger, A. P. [1957], "Economic Implications of Using Alternative Production Functions for Expressing Corn-Nitrogen Production Relationships," North Carol. Ag. Exp. Sta. Tech. Bull. \#126, Raleigh, North Carolina.

Suits, D. B. et al. [1978], "Spline Functions Fitted by Standard Regression Methods," Review of Economics and Statistics 60:132-139.

Swanson, E. R. [1963], "The Static Theory of the Firm and Three Laws of Plan Growth, " Soil Science 95(5):338-343.

Terman, G. L. and L. A. Nelson. [1976], "Comments on Use of Multiple Regression in Plant Analysis Interpretation," Agronomy Journal 68: 148-150.

Theil, H. [1971], Principles of Econometrics, John Wiley and Sons, Inc., New York.

Tisdale, S. L. and W. L. Nelson. [1975], Soil Fertility and Fertilizers, 3rd Ed., MacMillan, New York.

Tollini, H. and J. A. Seagraves. [1970], "Actual and Optimal Use of Fertilizer: The Case of Nitrogen on Corn in Eastern North Carolina," Economics Research Report \#14, North Carolina State University, Raleigh.

Tennessee Valley Authority [1976], "TVA Fertilizer Conference," Bull. Y-106.

Waggoner, P. E. and W. A. Norvell. [1979], "Fitting the Law of the Minimum to Fertilizer Applications and Crop Yields," Agronomy Journal 71:352-354.

White, H. [1980], "Using Least Squares to Approximate Unknown Regression Functions," International Economic Review 21(1):149-170.

White, K. J. [1978], "A General Computer Program for Econometric Methods - SHAZAM," Econometrica, January, 1968.

White, W. C. et al. [1958], "Evaluation of Residual Nitrogen in Soils," Agronomy Journal 50(1):255-259.

Wittwer, S. H. [1980], "Future Trends" in National Academy of Sciences: Lona Range Environmental Outlook, Proceedings of a Workshop Nov. 14-16, 1979, 80-85.


[^0]:    $\underline{1 /}^{1}$ For more about concepts of nutrient mobility, see for example Bray [1954].

[^1]:    1/Maximizing II may not be the farmers' objective under all circumstances. They may for instance maximize the expected utility of Il . This subject, however, transcends the scope of the present discussion.

[^2]:    ${ }^{1 /}$ The notation is the usual for the standard regression model $y=f(X)+\varepsilon$, $E$ being the symbol for the expectation operator.

[^3]:    ${ }^{1 /}$ For details of the derivation of the model and supporting arguments, reference should be made to Lanzer [1978].

[^4]:    $\underline{1}$ For other indirect techniques of measuring residual value of single applications of fertilizer, see for example Kennedy et al., [1973] and alternative approaches and extension in Helyar and Godden [1977].

[^5]:    $1 /$ The data set used in this study was kindly provided by Prof.
    S. A. Barber of the Department of Agronomy, Purdue University. Additional information regarding experimental procedures can be gleaned from the following publications: Barber, S.A. [1958], [1979], and [1980].

[^6]:    $1 /$ For details about the properties and some of the earlier applications of these two forms in crop response analysis, see for example Brown et a1. [1958] or any of the works of E. O. Heady cited in this study.

[^7]:    1/Private communication with Prof. Barber.

