

CORRELATED ERROR STRUCTURE IN PERENNIAL CROP EXPERIMENTS

by

Hólmgeir Björnsson and Foster Cady
Agricultural Research Institute, Reykjavík, Iceland
and
Biometrics Unit, Cornell University, Ithaca, N.Y. 14850

BU-469-M

July 1973

Abstract

Perennial crop experiments have the characteristic that the same plots are used during the several cycles of the experiment. Consequently the errors of observations during short intervals are more closely related than those of long intervals. A transformation of the yield data as explained by Battese, Fuller and Shrader (1972) can be carried out and an analysis on the nearly uncorrelated observations done. In the Box and Jenkins book, models are developed and procedures given for estimating appropriate covariance structures. Here various covariance structures useful in the analysis of perennial crop experiments are considered. A connectedness to split plot experiments and an analysis given by Patterson and Lowe (1970) are also discussed.

Hólmgæir Björnsson and Foster Cady

1. INTRODUCTION

All biological observations are subject to variation which is considered random and often called error if the sources of variation are not known or at least beyond control. One of the purposes of experimental design and statistical analysis is to provide a basis for estimating the contribution of error to the observations. A usual requirement in this connection is randomness of error. When the observations are taken in a determined order in space or time such as field plots arranged in a row, or annual plot harvests, errors are not independent but follow some structure such that errors of observations taken at short intervals in space or time are more closely related than those of long intervals. The use commonly made of this general knowledge in field experiments is to group plots located closely together into units called blocks, with the expectation of having smaller variation within blocks than in the field as a whole. In order to avoid systematic effects of fertility patterns remaining after blocking, treatments are allotted at random to plots within blocks.

Frequently, the nature of experiments is such that blocking and randomization cannot be used to eliminate systematic variation or dependence among errors. This is particularly true of observations ordered in time, often referred to as time series. Repeated observations on field plots are considered as a time series and so are the (annual) residuals after eliminating the annual design parameters (treatment and block effects).

In many situations the correlations between observations taken at different points in time are due to the influences of the earlier occurrences on the later;

this being regarded as a lag in the effect of the earlier occurrence. Hence, time intervals are called lags. Correlations between observations of constant lag usually stay constant; the correlations being called serial correlations or autocorrelations. These are generally found to decrease as lag increases.

1.1. Basic models for correlated residuals

Observations are recorded over a period of N years on m experimental units or plots where all plots receive the same annual treatment. An extension to the experimental situation with two or more treatments does not complicate the treatment of residuals.

The linear model can be simply written as

$$y_{it} = \mu_t + e_{it}, \quad (1.1)$$

where,

y_{it} is the observation on plot i ($i=1,2,\dots,m$) in
year t ($t=1,2,\dots,N$),

μ_t is the yield level in year t ,

and

e_{it} is a random error term.

Alternatively, the observation, y_{it} , is the outcome or realization at time t of the time series y_i of length N with the annual level μ_t and the random error e_{it} which is generated through a random time series process, also called a stochastic process.

The basic assumptions about the distribution of the error terms, e_{it} , are that the vectors of residuals, e_i , are samples from weakly stationary and independent

time series processes, identically distributed about their means, the means being random variables with expected value zero. In a second order or weakly stationary process the first and second moments depend only on time differences, i.e. they have constant means, and

$$\text{Cov}(e_{it}, e_{it-k}) = \text{Cov}(e_{i',t'}, e_{i',t'-k}) = \gamma_k \text{ for all } i, i' \text{ and } t, t', \quad (1.2)$$

where k is used to denote a constant difference in time or lag and γ_k is defined as the autocovariance of lag k , $\gamma_0 = \sigma^2$ is the variance of the series. The assumption of normality in addition to second order stationarity is sufficient to produce strict stationarity, i.e. all moments depend only on time differences (Box and Jenkins [1970, p. 30]). The other assumptions on the distribution of e_{it} can be summarized as

$$e_{it} \sim \text{IID}(0, \sigma^2), \gamma_0 = \sigma^2 \text{ for } t \text{ fixed and all } i \quad (1.3)$$

and

$$\text{Cov}(e_{it}, e_{i',t'}) = 0 \text{ for } i \neq i' \text{ and all } t, t' .$$

The autocovariance function of a time series process consists of all autocovariances of the process for $k \geq 0$.

Box and Jenkins (1970) estimate γ_k for a single series by

$$c_k = \frac{1}{N} \sum_{t=k+1}^N \hat{e}_{it} \hat{e}_{it-k} \quad (1.4)$$

where \hat{e} is an estimator of e . c_k is biased for $k \geq 1$ and the values of c_k are highly correlated with each other.

Corresponding to the autocovariance of lag k is the autocorrelation coefficient of lag k defined by

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (1.5)$$

A variety of special cases of the covariance structure (γ_k) will be considered:

(a) The simplest autocovariance function is to let

$$\gamma_k = 0 \quad \text{for } k > 0 \quad .$$

In this case the errors are generated by a completely random process and are identically and independently distributed with mean zero and variance σ^2 , i.e.,

$$e_{it} \sim \text{IID}(0, \sigma^2) \quad . \quad (1.6)$$

Univariate ANOVA methods are applicable for this error model.

(b) Another simple autocovariance function is when the covariances are all equal and nonzero. For the moment, only positive values for γ_k will be considered:

$$\sigma^2 > \gamma_k = b > 0 \quad \text{for } k > 0 \quad .$$

The case with $b < 0$ is introduced in 1.2. The error e_{it} can be split into two independent variance components, $\gamma_k = b$ being the component common to all observations on one plot, and $\sigma^2 - b$ being a component uncorrelated among annual observations. This variance structure can be represented by the model

$$e_{it} = w_i + a_{it} \quad (1.7)$$

where w_i and a_{it} are random variables

$$w_i \sim \text{IID}(0, \sigma_w^2) ; \quad \sigma_w^2 = b ,$$

$$a_{it} \sim \text{IID}(0, \sigma_a^2) ; \quad \sigma_a^2 = \sigma^2 - b \quad .$$

This is the split-plot error structure, also treated in 1.2. The term w_i which is common to all observations on the plot is called plot error in field experiments (Patterson, 1953). The inclusion of plot error as a component of the

autocovariance function in (1.7) and later on differs from applied time series analysis where autocovariances are calculated around the mean of the series.

A time series with trend is nonstationary. Linear trend can be eliminated by taking differences which also eliminates the mean. After differencing the series may be stationary. This practice is comparable to experimental situations where the split-plot model (1.7) is applicable and comparisons are made on sub-plots within whole plots. In this case differences of the kind

$$\begin{aligned}\Delta e_t &= e_{it} - e_{it'} \\ &= w_i - a_{it} - (w_i - a_{it'}) \\ &= a_{it} - a_{it'}\end{aligned}$$

are taken which eliminate the plot error w_i .

When, on the other hand, comparisons are to be made on a between-plot basis, the plot error or random mean of the series enters, a feature which distinguishes the experimental situation from time series analysis.

(c) If γ_1 is larger than the covariances with longer lags and if $b = 0$, i.e.

$$\sigma^2 > \gamma_1 > \gamma_k = 0, \text{ for } k > 1,$$

the model for e_{it} corresponding to this autocovariance function is usually written

$$e_{it} = Z_{it} = a_{it} - \theta_1 a_{it-1} \tag{1.8}$$

where

$$a_{it} \sim \text{IID}(0, \sigma_a^2)$$

This error structure also holds for negative values of γ_1 , i.e. θ_1 can be both positive and negative. This is a model for residual effects and is generally referred to as the first order moving average process, MA(1), with the parameters θ_1 . Keeping in mind the independence of a_{it} the covariance of lag 1 is

$$\begin{aligned}
 \gamma_1 &= \text{Cov}(Z_{it}, Z_{it-1}) \\
 &= \text{Cov}(a_{it} - \theta_1 a_{it-1}, a_{it-1} - \theta_1 a_{it-2}) \\
 &= \text{Cov}(-\theta_1 a_{it-1}, a_{it-1}) \\
 &= -\theta_1 \sigma_a^2,
 \end{aligned} \tag{1.9}$$

and the variance is

$$\begin{aligned}
 \gamma_0 &= \sigma_Z^2 = E(Z_{it}^2) = E(a_{it} - \theta_1 a_{it-1})^2 \\
 &= (1 + \theta_1^2) \sigma_a^2.
 \end{aligned}$$

The relationship between ρ_1 and θ_1 is

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta_1}{1 + \theta_1^2} \tag{1.10}$$

The first order moving average parameter with ρ_1 known can be obtained as a solution to the second degree polynomial

$$\theta_1^2 + \frac{\theta_1}{\rho_1} + 1 = 0,$$

which has two roots with the product 1, of which the root with $|\theta_1| < 1$ is selected.

(d) The situation is the same as before except that $b > 0$

$$\sigma^2 > \gamma_1 > \gamma_k = b > 0, \text{ for } k > 1.$$

This corresponds to the error model

$$\begin{aligned}
 e_{it} &= w_i + Z_{it} \\
 &= w_i + a_{it} - \theta_1 a_{it-1}
 \end{aligned} \tag{1.11}$$

where, as in (1.7), w_i is used to denote the random mean of the series, $\text{Var}(w_i) = \sigma_w^2$, Z_{it} is used to denote the stochastic process of deviations from the mean, $\text{Var}(Z_{it})$

$= \sigma_Z^2$ and, as in (1.7) and (1.8), a_{it} is used to denote the random contribution to e_{it} which is independent of previous outcomes, $\text{Var}(a_{it}) = \sigma_a^2$. However, as in (1.8) and (1.11), the a_{it} do influence future outcomes. The term a_{it} is often referred to as random shock or white noise.

Using these definitions, $e_{it} = Z_{it} = a_{it}$ in (1.6) and $e_{it} = Z_{it}$ in (1.7).

It follows from (1.7) and (1.9) that

$$\begin{aligned}\sigma^2 &= \gamma_0 = \sigma_w^2 + \sigma_Z^2 \\ &= \sigma_w^2 + (1 + \theta_1^2)\sigma_a^2, \\ \gamma_1 &= \sigma_w^2 - \theta_1\sigma_a^2, \\ \gamma_k &= \sigma_w^2 = b \quad \text{for } k > 1\end{aligned}$$

and

$$\rho_1 = \frac{\sigma_w^2 - \theta_1\sigma_a^2}{\sigma_w^2 + (1 + \theta_1^2)\sigma_a^2}$$

which leads to the second degree polynomial in θ_1

$$\theta_1^2 + \frac{\theta_1}{\rho_1} + 1 - \frac{1 - \rho_1}{\rho_1} \frac{\sigma_w^2}{\sigma_a^2} = 0 \quad . \quad (1.12)$$

When attention is limited to comparisons within series, only the moving average component of (1.11) deserves attention.

(e) Let all autocovariances be positive but approaching zero as lag increases,

$$\gamma_0 > \gamma_1 > \dots, \gamma_k \rightarrow 0 \text{ for } k \text{ large} .$$

If in addition

$$\frac{\gamma_k}{\gamma_{k-1}} = \theta_1 = \rho_1 \text{ for } k \geq 1$$

then the error model is

$$\begin{aligned} e_{it} = Z_{it} &= \phi_1 Z_{it-1} + a_{it} , \\ \text{Var}(Z_{it}) &= \sigma_Z^2 = \phi_1^2 \sigma_Z^2 + \sigma_a^2 . \end{aligned} \quad (1.13)$$

This expression has the appearance of a regression equation with Z_{it} being dependent on the independent next preceding observation Z_{it-1} of the same variable and, hence, it is called autoregressive with the autoregressive coefficient ϕ_1 . The intercept of the autoregressive equation is the random shock occurrence a_{it} . When only the immediately preceding observation is involved in the process it is called first order autoregressive AR(1). An autoregressive process is constantly changing levels by the random shock a_{it} , but their effects die out.

Since

$$Z_{it-1} = \phi_1 Z_{it-2} + a_{it-1}$$

then

$$\begin{aligned} Z_{it} &= \phi_1^2 Z_{it-2} + \phi_1 a_{it-1} + a_{it} \\ &= \phi_1^k Z_{it-k} + \sum_{h=0}^{k-1} \phi_1^h a_{it-h} . \end{aligned} \quad (1.14)$$

Given the independence of a_{it} with Z_{it-h} , $h > 0$ it follows that

$$\begin{aligned} \gamma_k &= \text{Cov}(Z_{it}, Z_{it-k}) \\ &= \text{Cov}\left(\phi_1^k Z_{it-k} + \sum_{h=0}^{k-1} \phi_1^h a_{it-h}, Z_{it-k}\right) \\ &= \phi_1^k \sigma_Z^2 \end{aligned} \quad (1.15)$$

and

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi_1^k . \quad (1.16)$$

The relationship between the parameters follows from (1.13)

$$\sigma_Z^2 = \frac{\sigma_a^2}{1 - \phi_1^2} . \quad (1.17)$$

Negative values of ϕ_1 will cause the autocovariance function to alternate in signs as it approaches zero. Since σ_Z^2 is positive, $-1 < \phi_1 < 1$.

(f) Finally, let the situation be that of (e) except that at large lags the covariances approach a value $b > 0$:

$$Y_0 > Y_1 > \dots, Y_k \rightarrow b > 0 \text{ for } k \text{ large}$$

and

$$\frac{Y_k - b}{Y_{k-1} - b} = \phi_1 < \phi_1 \text{ for } k \geq 1 ,$$

then the error model is

$$\begin{aligned} e_{it} &= w_i + Z_{it} & (1.18) \\ &= w_i + \phi_1 Z_{it-1} + a_{it} \\ \text{Var}(e_{it}) &= \sigma^2 = \sigma_w^2 + \sigma_Z^2 \\ &= \sigma_w^2 + \frac{\sigma_a^2}{1 - \phi_1^2} \end{aligned}$$

and

$$Y_k = \sigma_w^2 + \phi_1^k \sigma_Z^2 .$$

The situation (f) as stated is limited to $\phi_1 > 0$. With $\phi_1 < 0$ the autocovariance function will fluctuate as it approaches b with increasing lag.

1.2. Comparison of the AR(1) and split-plot models

Of the situations (a) to (f) in part 1.1 the ones of greatest interest are (b) the split-plot model, and (e) the first order autoregressive. As given in equation (1.6) the split-plot model is composed of two variance components σ_w^2 and σ_a^2 . Expected values of the whole plot error E_a and the sub-plot error E_b in analysis of variance are

$$E(E_a) = \sigma_a^2 + N\sigma_w^2$$

and

$$E(E_b) = \sigma_a^2 \quad (1.19)$$

These expected values and model (1.7) are valid only for positive values of γ_k . The error model can be extended to negative values of γ_k by introducing the competition effect (Federer, 1955, p. 59) or, which is preferred here, by introducing the plot correlation or intraclass correlation

$$\rho = \frac{\gamma_k}{\gamma_0} \quad (1.20)$$

which is constant for constant γ_k , but can take on both positive and negative values. Expected values of the whole plot error E_a and the sub-plot error E_b are (see, e.g., Cochran and Cox [1957, p. 295])

$$E(E_a) = \sigma^2(1 + [N - 1]\rho)$$

$$E(E_b) = \sigma^2(1 - \rho) \quad (1.20)$$

In matrix notation the split-plot error model, which also is referred to as the uniform error model, is

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho & \dots & \dots \\ \rho & 1 & \rho & & . \\ \rho & \rho & 1 & & . \\ . & & & . & . \\ . & & & & . \\ . & & & & . \\ \rho & \dots & \dots & \dots & 1 \end{bmatrix} = \sigma^2 (1 - \rho)I + \rho J, \quad (1.22)$$

and the first order autoregressive model is

$$\sigma_Z^2 \begin{bmatrix} 1 & \rho_1 & \rho_1^2 & \dots & \rho_1^{N-1} \\ \rho_1 & 1 & \rho_1 & & . \\ \rho_1^2 & \rho_1 & 1 & & . \\ . & & & . & . \\ . & & & & . \\ . & & & & . \\ \rho_1^{N-1} & \dots & \rho_1 & & 1 \end{bmatrix} = \sigma_Z^2 R_{\phi_1}, \quad (1.23)$$

where R_{ϕ_1} is used to denote the correlation matrix generated by the first order autoregressive coefficient $\phi_1 = \rho_1$.

This comparison shows that the problem of estimating error in both the first order autoregressive model and the split-plot model can be regarded as an estimation of a variance and a correlation coefficient. However, when the error matrix is used for testing hypotheses about treatments it is the estimation of composite error terms (like E_a) which is ultimately of interest. In the split-plot case the error variances can always be constructed as linear combinations of two independently estimated errors, E_a and E_b , but with the autocorrelation error model this cannot be done.

1.3. The error model of Patterson and Lowe (1970)

Patterson and Lowe (1970) incorporated the general observation that plot correlations (autocorrelations) do decrease with time by replacing the constant plot error w_i in the split-plot error model (1.7) by an autoregressive component of the type (1.13). The independent error or noise was thus divided in two components; one, the shock which operates on the autoregressive component and two, the error of observation which is not correlated with any other error:

$$\begin{aligned} e_{it} &= Z_{it} = Z_{it}^! + a_{it}^{!!} \\ &= \phi_1^! Z_{it-1}^! + a_{it}^! + a_{it}^{!!} \\ &= \phi_1^! Z_{it-1}^! + a_{it} \end{aligned} \tag{1.24}$$

where Z_{it} , the outcome of the random process, is split in $Z_{it}^!$, the autoregressive component and $a_{it}^{!!}$, the independent random error of observation. $\phi_1^!$ is used to denote a first order autoregressive parameter when operating on a component $Z_{it}^!$ of the random process Z_{it} . $a_{it}^!$ is the random shock operating on the autoregressive process and $a_{it} = a_{it}^! + a_{it}^{!!}$ is the white noise component of Z_{it} which is independent of earlier errors.

$$\begin{aligned} \text{Var}(Z_{it}) &= \sigma_Z^2 \\ &= \text{Var}(Z_{it}^!) + \text{Var}(a_{it}^{!!}) = \sigma_{Z^!}^2 + \sigma_{a^{!!}}^2 \end{aligned}$$

$$\text{Var}(a_{it}^!) = \sigma_{a^!}^2$$

$$\text{Var}(a_{it}) = \sigma_a^2 = \sigma_{a^!}^2 + \sigma_{a^{!!}}^2$$

The process (1.24) is a first order autoregressive process with mean zero and the independent error of observation (white noise) added. It is shown by Box and Jenkins (1970, p. 122) that the addition of white noise to an autoregressive model can be formulated by including a moving average term and letting the autoregressive parameter operate on Z_{it} rather than $Z_{it}^!$:

$$Z_{it} = \phi_1 Z_{it-1} + a_{it} - \theta_1 a_{t-1} \quad (1.25)$$

This mixed first order autoregressive, first order moving average process is commonly denoted ARMA(1,1) and is one of the most important time series models.

The following relationship between parameters can be obtained from the two expressions (1.24) and (1.25) of ARMA(1,1) model:

$$\gamma_1 = \phi_1' \sigma_{Z'}^2 = \phi_1 \sigma_Z^2 \quad .$$

Then,

$$\frac{\phi_1}{\phi_1'} = \frac{\sigma_{Z'}^2}{\sigma_{Z'}^2 + \sigma_{a'}^2} = \frac{\sigma_Z^2 - \sigma_{a'}^2}{\sigma_Z^2} \quad (1.26a)$$

and for $k > 1$

$$\gamma_k = \phi_1^{k-1} \rho_1 \sigma_Z^2 \quad .$$

The ARMA(1,1) would have fitted into the treatment as case (g) with the following statement about the relationship between the covariances: The situation is the same as in (e) except that $\phi_1 > \rho_1$:

$$\gamma_0 > \gamma_1 > \dots > \gamma_k \rightarrow 0 \text{ for } k \text{ large}$$

$$\frac{\gamma_k}{\gamma_{k-1}} = \phi_1 \text{ for } k \geq 2$$

$$\frac{\gamma_1}{\gamma_0} = \rho_1 < \phi_1$$

This situation as described limits both ϕ_1 and θ_1 to positive values; but in general an ARMA(1,1) process may have negative parameter values.

In the example provided by Patterson and Lowe (1970) observations were taken at 4-year intervals, and lag 1 is therefore taken to mean an interval at 4 years.

The estimate of ρ_1 was 0.2 and estimates of ϕ_1' lay between 0.8 and 1.0. If $\phi_1' = 0.8$, $\rho_1 = 0.2$ and $\sigma_Z^2 = 1$, then $\sigma_{Z'}^2 = \frac{0.2}{0.8} = 0.25$, $\sigma_{a'}^2 = 1 - 0.25 = 0.75$, $\sigma_{a'}^2 = 0.25(1 - 0.8^2) = 0.09$, $\sigma_a^2 = 0.09 + 0.75 = 0.84$, $\phi_1 = \frac{0.25}{0.25 + 0.75} = 0.25$ and $\theta_1 = 0.052$.

The authors gave examples of how the assumption of uniformity leads to biased estimates of error of linear trend when in fact model (1.24) is valid and $\rho = 0.2$. With 6 periods and $\phi_1' = 0.8$ the error would be underestimated by 15.6% and with $\phi_1' = 0.5$ by 22.9%. These results demonstrate that the autocorrelation structure of experimental errors cannot always be neglected.

1.4. Extensions to more complex situations

For short-term experiments or in rotation experiments with crops returning at intervals of a few years the models considered so far will most often be sufficient. For long-term experiments this may not be the case.

Compared to the unrestricted variance-covariance matrix Σ of Morrison (1967), the reduction of unknown parameters from $N(N - 1)/2$ to two as in (1.7) and (1.13) or three as in (1.18) and (1.25) should lead to a great increase in power, but the strong assumptions involved may not hold. The most expected violation of the assumptions would be inequality of the variances σ_a^2 which in model (1.25) could apply to both components, $\sigma_{a'}^2$ and $\sigma_{a''}^2$.

The first order autoregressive coefficient ϕ_1 may be variable in time. Severe year effects which reduce stand on plots with permanent grass would be interpreted as larger than usual shocks. In other words the annual value for σ_a^2 is larger than usual. This could lead to increased variance of observation, σ_Z^2 , or be counteracted by a reduced value for ϕ_1 . Years after and before this particular year would be less correlated than expected. Allowing ϕ_1 to take on new value every year changes expression (1.15) to

$$\text{Cov}[Z_{it}, Z_{it'}] = \sigma_Z^2 \prod_{h=t}^{t'} \phi_{1h} \quad (1.26b)$$

for $t < t'$

This covariance structure also arises when the time intervals between observations are variable. Morrison (1967, p.296) indicated the use of this model for estimating unknown time intervals between observations. People often express their subjective feeling of a hard winter by calling it a long winter. It is conjectured that the severity of winter might be measured on a scale with

$$\bar{\phi}_1 = \left\{ \prod_{h=1}^N \phi_{1h} \right\}^{1/N} \quad (1.27)$$

as a unit.

Situations similar to those exemplified above, although more difficult to visualize, are likely to be represented by a second order autoregressive process, AR(2):

$$Z_{it} = a_{it} + \phi_1 A_{it-1} + \phi_2 Z_{it-2} \quad (1.28)$$

This process belongs to the general class of autoregressive moving average processes of order p and q, ARMA(p,q) (Box and Jenkins [1970]).

1.5. The role of experimental designs

It has been assumed that the errors are independent between plots. As mentioned in the introductory remarks, however, the residuals are expected to follow a pattern in the field and independence is secured by allotting treatments at random to the plots. If the errors follow model (1.18) and the a_{it} are independent between plots, only the plot error w_1 is available for randomization and can be reduced by blocking. The use of an efficient design is of particular importance

when the experiment is going to be continued for a long period. However, even the a_{it} may follow patterns within years. Drainage gradient within an experiment will give different relationships in dry and humid years, which means that the a_{it} are more closely related on plots with similar drainage levels. This example can be extended to the situation that the drainage level goes through irreversible changes in time, causing irreversible changes in fertility patterns which would enter as nonstationarity into the time series model. Complications of the kind exemplified can arise for other reasons, but will be greatly reduced by the use of efficient plot arrangements: size, shape and blocking.

References

- Box, G. E. P. and G. M. Jenkins. 1970. Time Series Analysis, Forecasting, and Control. Holden-Day Inc., San Francisco.
- Federer, W. T. 1955. Experimental Design. The MacMillan Company, New York.
- Morrison, D. F. 1967. Multivariate Statistical Methods. McGraw-Hill, New York.
- Patterson, H. D. and B. I. Lowe. 1970. The errors of long term experiments. J. Agr. Sci. 74: 53-60.
- Battese, G. E., W. A. Fuller and W. D. Shrader. 1972. Analysis of crop-rotation experiments, with application to the Iowa Carrington-Clyde rotation-fertility experiments. Agriculture Experiment Station Research Bulletin 574, Iowa State University.