Spatial Analysis and Efficiency of Systematic Designs in Intercropping Experiments

SYMON PETER WANDIEMBE

Submitted in partial fulfilment of the academic

requirements for the degree of

MASTER OF SCIENCE

in

Statistics

in the

School of Mathematics, Statistics and

Information Technology,

University of Natal

Pietermaritzburg

December, 2002.

Dedication

To

Auntie Grace, Irene, Doreen and Lilly.

The works of His hands are faithful and just;
all His precepts are trustworthy.

PSALM 111: 7.

Declaration

The research work described in this thesis was carried out in the School of Mathematics, Statistics and Information Technology, University of Natal, Pietermaritzburg, from January 2001 to December 2002, under supervision of Dr. Peter M. Njuho.

The work represents original work by the author and have not otherwise been submitted in any form for any degree or diploma to any University. Where use has been made of the work of others it is duly acknowledged.

Student:..... Date: 01-04-03

Symon P. Wandiembe

Supervisor: ..... Date: 01-04-2003

Dr. Peter M. Njuho

Acknowledgement

I am very much indebted to my supervisor, Dr. Peter Njuho for his guidance and support throughout this study. My deep appreciation goes to Professor Linda Haines and Mr. Harvey Dicks for their invaluable advice and suggestions. I owe a great debt of gratitude to the staff members and postgraduate students at School of Mathematics, Statistics and Information Technology, UNP for the hospitable environment they provided during my study.

It is a pleasure to acknowledge the encouragement provided by friends especially Irene, Maureen, Thabile, Thomas, Robert and Chris. My special thanks is also due my auntie's family.

I also with deep appreciation thank Rockefeller Foundation for funding my studies at the University of Natal. In the same way, I sincerely thank the collaborators at Makerere University, Uganda especially Professor David Osiru. I would like to thank Professor Adipala Ekwamu and Dr. Margaret Nabasirye for their continual support during my studies.

I thank everyone though not mentioned by name but have greatly contributed to my academic success.

To all, may God bless you.

Abstract

In studies involving intercropping plant populations, the main interest is to locate the position of the maximum response or to study the response pattern. Such studies normally require many plant population levels. Thus, designs such as spacing systematic designs that minimise experimental land area are desired. Randomised block designs may not perform well as they allow few population levels which may not span the maximum or enable exploration of other features of the response surface. However, lack of complete randomisation in systematic designs may imply spatial variability (large-scale and small-scale variations i.e. trend and spatial dependence) in observations. There is no correct statistical method laid out for data analysis from such designs. Given that spacing systematic designs are not well explored in literature, the main thrusts of this study are two fold; namely, to explore the use of spatial modelling techniques in analysing and modelling data from systematic designs, and to evaluate the efficiency of systematic designs used in intercropping experiments. Three classes of models for trend and error modelling are explored/introduced. These include spatial linear mixed models, semi-parametric mixed models and beta-hat models incorporating spatial variability. The reliability and precision of these methods are demonstrated. Relative efficiency of systematic designs to completely randomised design are evaluated. The analysis of data from systematic designs is shown be easily implemented. Measures of efficiency that include ϕ_p directed measures (A and E criteria), D_1 and D^s efficiencies for regression parameters, and power are used. Systematic designs are shown to be efficient; on average 72% for A and E- efficiencies and 93% for D_1 and D^s efficiencies. Overall, these results suggest that systematic designs are suitable and reliable for intercropping plant population studies.

Contents

List of Tables

Chapter 1

Introduction

1.1 Background

Intercropping is an 'old-age' farming practice that refers to the growing of two or more crops simultaneously on the same piece of land. During the past two decades or so considerable research attention has been directed to understanding the possible intercropping advantages (Osiru and Willey, 1972; Mead and Stern, 1980; Ocaya, 1998). These advantages are measured in terms of responses such as crop yield, land equivalent ratios, economic returns, coefficients of aggressivity, calorie value, etc. Despite these substantial research efforts, little attention has been given to experimental designs. All too often experimenters and statisticians feel that their choice of an experimental design is limited to those appearing in the tables or in the literature such as, for example those in Cochran and Cox (1964). The experiment should be considered as it is to be conducted, rather than being changed to fit a design published in the literature or in the tables. Direct extensions of procedures and concepts for sole cropping to intercropping are often inappropriate and can sometimes be misleading (Federer, 1993).

The need for designing experiments specifically for intercropping was first outlined

by Willey (1979). Subsequently, many other agronomists and statisticians have similarly expressed their concern about the inadequate consideration of the designs for intercropping experiments (Mead and Stern, 1980; Mead and Rilley, 1981; Finney, 1988; Federer, 1993; Peterson, 1994). There are indeed many aspects of intercropping experiments that require careful considerations. However, the most pressing one is in the area of plant populations i.e. plant density and spatial arrangements. Research in these factors (plant density and spatial arrangements) has gone on for a long time, yet their individual effects have seldom been distinguished. Insufficient identification of the various relations involved and little development of experimental designs to examine them could be the contributing factors (Willey and Rao, 1981). Establishing the pattern of response and identifying the optimal plant density combinations at early stages of intercropping experiments, and also distinguishing their effects is very important. These two factors define both species inter-competition and intra-competition, which in turn determines the intercrop yields.

The position of the maximum response in an intercrop can vary considerably from that predictable from monoculture (Federer, 1993). Determining such patterns based on conventional randomised block factorial experiments has some limitations. Since conventional randomised block designs do not allow very many treatment levels there is a risk that the treatment levels so included do not span the optimum or that if they do, they are so far apart that some vital feature of the response surface is not detected. Some researchers have employed systematic designs suggested by Nelder (1962) to overcome these problems (Willey and Lakhani, 1976; Huxley and Maingu, 1978; Wahua and Miller, 1978; Mead and Riley, 1981). These designs allow many levels of plant populations and use a minimum space of land compared to randomised designs since they do not need guard rows. They thus for the same land area accommodate more treatment levels than randomised designs.

Chapter 2

Review of Intercropping Experiments

2.1 General overview

2.1.1 Introduction

The importance of plant populations in relation to the intercrop yields is discussed. The evaluation of intercropping advantages and some experimental designs that are employed in intercropping studies are briefly discussed. By definition, intercropping is the growing of two or more crops simultaneously or sequentially on the same piece of land (Federer, 1993). The practice is used extensively in tropical agriculture, and will no doubt become important in temperate zone agriculture (Wahua and Miller, 1978; Federer, 1993). Mixed cropping is one of the least expensive methods of increasing productivity of crop fields with limited resource capacity.

Intercropping is a space-dependent form of multiple cropping. Thus, as already mentioned, when two or more crops are grown together, each must have adequate space to maximise co-operation within and minimise competition between them. The

following factors are necessary to consider when conducting an intercrop experiment: spatial arrangement; plant density; maturity dates of crops being grown together and plant architecture, and management. Spatial arrangement greatly contributes to the performance of crops as it affects edaphic interactions and determines light penetration into the canopies of both the taller and shorter components of an intercrop.

2.1.2 Plant population and spatial arrangement

Plant population refers to the total number of plants in an area. The plant population per given area is called plant density. Spatial arrangement in intercropping studies refers to both the relative positions on the ground of one plant to another within the same species (rectangularity) and the relative arrangement of plants of one species to another. Plant populations and spatial arrangement are factors that are commonly studied in intercropping research (Willey and Rao, 1981). However, as mentioned in Chapter 1 their individual effects have seldom been distinguished because there has been insufficient investigation of the various relations involved and little development in experimental designs to examine them (Willey and Rao, 1981).

Research efforts on the relationship between plant density or populations and crop yield in monoculture has been widely studied (Nelder, 1962; Willey and Heath, 1969). Generally, yield y_{l_i} $(i = 1, 2, \ldots, n)$ can be modelled as a function of density, $f(\cdot)$;

$$y_{l_i} = f(d_{l_i}) + \varepsilon_{l_i} \tag{2.1}$$

where $f(d_{l_i})$ can be a polynomial, linear or quadratic in density levels (d_{l_i}), or an inverse quadratic polynomial, and ε_{l_i} are random errors. To extend the concept of the yield-density model to intercropping, Wijesinha (1981) and Federer (1993) simply defined an additive effect/interaction effect $\gamma_{i(j)}(l_i, l_j)$ of density level l_j of crop j grown with density level l_i of crop i. That is, additive effect $\gamma_{i(j)}(l_i, l_j)$ is assumed to be the function

of densities d_{il_i} and d_{jl_j}, where d_{il_i} is the density level l_i of crop i and d_{jl_j} is the density level l_j of crop j. Using the Federer (1993) notation,

$$y_{i(j)l_il_j} = f(d_{il_i}) + \gamma_{i(j)}(d_{il_i}, d_{jl_j}) + \varepsilon_{i(j)l_il_j} \tag{2.2}$$

where $l_i = 0, 1, 2, \ldots, a$ (a density levels of crop i) and $l_j = 0, 1, 2, \ldots, b$ (b density levels of crop j); $l_i \neq l_j$ and where $y_{i(j)l_il_j}$ is the yield of crop i grown at plant density level l_i with plant density level l_j of crop j. $f(d_{il_i})$ is the function of density of crop i at level l_i while $\gamma_{i(j)}(d_{il_i}, d_{jl_j})$ gives the additive effect, $\gamma_{i(j)}(l_i, l_j)$, of crop j at density level l_j on yield of crop i. $\varepsilon_{i(j)}(l_i, l_j)$ is the random component of variation with variance σ_i^2 for each individual observation of crop i. The covariance between the random components of variation of individual yields of the intercrop of crop i and j is σ_{ij}. The variance structure of $\gamma_{i(j)}(l_i, l_j)$ have been given in detail by Wijesinha (Federer, 1993).

In this work, the linear response model is adopted because it is easily lends itself to the well-developed regression techniques. Thus

$$y_{i(j)l_il_jh} = \beta_{0i} + \beta_{1i}d_{il_i} + \gamma_{i(j)}(d_{il_i}, d_{jl_j}) + \varepsilon_{i(j)l_il_jh} \tag{2.3}$$

where $y_{i(j)l_il_jh}$ is the yield of the hth ($h = 1, 2, \ldots, r$) replicate for crop i grown at plant density level l_i with plant density level l_j of crop j. The random components of variation, $\varepsilon_{i(j)l_il_jh}$, are assumed independent and identically normally distributed with mean zero and common variance σ_i^2 i.e. $\varepsilon_{i(j)l_il_j} \sim iidN(0, \sigma_i^2)$, β_{0i} and β_{1i} are regression parameters for intercept and slope respectively. The magnitudes of the parameters indicate the importance of change in planting density, as well as the effect of intercrop and its planting density on the crop being evaluated. A pattern in the $\gamma_{i(j)}(l_i, l_j)$'s or $\gamma_{j(i)}(l_j, l_i)$'s would indicate the functional relationship due to planting densities of two crops.

Apart from the above developments a general linear statistical model can be used to model yield as a function of density. For example, the plant densities can be treated as

variates or factors and a factorial analysis of variance can be performed. Further details are provided in Section 2.2.

The determination of the effect of the range of planting densities, spatial arrangements and other factors when two crops are grown together entails frequent use of measures of advantages of intercropping. Some of these measures are outlined in the following sections.

2.1.3 Some measures of intercropping advantages

The main purpose of this section is to present the review of derived responses of intercropping studies in the context of plant population studies. These derived responses and other directly observed responses provide the measures of intercropping advantages over monoculture. In defining the advantages of intercropping, one needs to first define whether the two crops grown together are both main crops or one is a supplemental crop. The inclusion of a supplemental crop in a cropping system is partly to check the effect it may have on the main crop, and thus, analysis differs slightly.

There are various indices for examining the benefits of intercrops. Later in this chapter, in Section 2.2, modelling of these responses is suggested. The commonly used indices/derived responses include among others the following.

Relative yield totals (RYT)

This was suggested by de Wit and Van den Bergh (1965) who were mainly interested in replacement series competition experiments.

$$RYT = r_i + r_j$$

where

$$r_i = \frac{\bar{y}_{i(j)l_i..m}}{\bar{y}_{i.s}} \text{ and } r_j = \frac{\bar{y}_{j(i)l_j..m}}{\bar{y}_{j.s}}$$

where $\bar{y}_{i(j)l_i..m}$ is the average yield of density level l_i of crop species i averaged over various density levels (l_j) of crop species j and r replicates, and $\bar{y}_{i.s}$ is the sole crop yields of crop species i averaged over r replicates. The terms $\bar{y}_{j(i)l_i..m}$ and $\bar{y}_{j.s}$ are defined similarly for crop species j. When the value of $RYT > 1$ then intercropping offers an advantage over sole cropping otherwise no advantage is gained.

Land equivalent ratio (LER)

This is a measure of relative land area required to produce the same yields by sole cropping as those achieved by intercropping (Willey and Osiru, 1972). An LER of more than 1 means an intercropping advantage. For instance, an LER of 1.3 indicates a yield advantage of 30% (in other words, 30% of more land would be required for sole crops to produce the same yield as intercropping). It represents the increased biological efficiency achieved by two crops grown together (Mead and Willey, 1980). It is defined by Willey and Osiru (1972) as

$$LER = \frac{Y_{mi}}{y_i} + \frac{Y_{mj}}{y_j} = L_i + L_j$$

where L_i and L_j are *partial* LERs of crop i and j respectively. For a given land area, Y_{mi} is the yield of crop i intercropped with crop j and Y_{mj} is the yield of crop j intercropped with crop i, and y_i and y_j are sole crop yields of crops i and j, respectively. Mead and Willey (1980) referred to y_i and y_j as *standardising* factors. There are various forms of standardizing intercrop yields (see for example, Federer, 1993). The form of standardization should vary according to the objective of the experiment (Mead and Willey, 1980). In this work, the following form of LER will be used.

$$L^{+}_{hijl_il_j} = \frac{y_{i(j)l_il_jh}}{\bar{y}_{i.}} + \frac{y_{j(i)l_il_jh}}{\bar{y}_{j.}} \tag{2.4}$$

where $L^{+}_{hijl_il_j}$ is hth replicate total LER computed at intercrop of density levels l_i and l_j and is thus a combination of partial LERs $\left(\frac{y_{i(j)l_il_jh}}{\bar{y}_{i.}}\right)$ for crop *1* and $\left(\frac{y_{j(i)l_il_jh}}{\bar{y}_{j.}}\right)$ for crop

j. where $\bar{y}_{i.}$ and $\bar{y}_{j.}$ are the optimum yields of crop i and j respectively, when grown at their optimum population and spacing averaged over farms or years or experiments, etc. Huxley and Maingu (1978) suggested that LER be calculated using the *optimum pure stand yield* for each component especially when the study involves plant populations and spacings. The LER calculated using optimum pure stand yield for each component avoids the confounding of beneficial interactions between component crops with response to changes in plant population. Thus, a farmer can ascertain whether he is technically better off with mixtures or sole crops. Federer (1993), indicates that LER calculated using optimum pure stand yield are not prone to high correlations with each other. The LER generated by this method has a distribution which does not deviate much from normal distribution (Federer and Schwager, 1982).

Many extensions to LER exist and they include relative LER, staple LER, effective LER, etc. (Mead and Riley, 1981). For example, of great importance for statistical analysis is the relative LER which is given by

$$L^{*}_{hijl_il_j} = y_{i(j)l_il_jh} + \frac{\bar{y}_{i.}}{\bar{y}_{j.}} y_{j(i)l_il_jh} \tag{2.5}$$

Relative LER, in general, can be taken as a linear combination of crop responses.

$$LER = K_1 y_{i(j)l_il_jh} + K_2 y_{j(i)l_il_jh} \tag{2.6}$$

The coefficients K_1 and K_2 could be crop values; protein, carbohydrate conversion or calorie values, prices, farmers' values, or coefficients obtained from multivariate analysis, etc. (Balaam, 1986).

Area-time equivalent ratio (ATER)

Area-Time Equivalent Ratio takes into account the time the crop occupies the land from sowing to harvesting. This method permits an evaluation of crops on a yield-per-day

basis.

$$ATER = \frac{L_i T_1 + L_j T_2}{T}$$

where L_i and L_j are partial LERs of crop i and j respectively while T_1 and T_2 are their respective durations. The term $T\,(T_1 + T_2 \leq T)$ is the total duration of the intercrop. An $ATER > 1$ indicates an advantage of intercropping.

Indices of competition

One of indices of competition is the Competitive Ratio (CR). It is computed as

$$CR = \left(\frac{\bar{y}_{i(j)l_i..m}}{\bar{y}_{i.s}\lambda_1}\right) \Big/ \left(\frac{\bar{y}_{j(i)l_j..m}}{\bar{y}_{j.s}\lambda_2}\right) = \left(\frac{L_i}{L_j}\right).\left(\frac{\lambda_2}{\lambda_1}\right)$$

where λ_1 and λ_2 are the sown proportions of crop species i and j respectively. Another index is the *Relative Crowding Coefficient, K*. A coefficient $K = K_1 K_2$, where K_i is defined for each species, where $K_1 = \bar{y}_{i(j)l_i..m}\lambda_1\lambda_2/(\bar{y}_{i.s} - \bar{y}_{i(j)l_i..m})\lambda_1$, and $K_2 = \bar{y}_{i(j)l_j..m}\lambda_1/(\bar{y}_{j.s} - \bar{y}_{j(i)l_j..m})\lambda_2$ where $\bar{y}_{i(j)l_i..m}$ and $\bar{y}_{j(i)l_j..m}$ are the intercrop yields of respective crops i and j, $\bar{y}_{i.s}$ and $\bar{y}_{j.s}$ are the respective sole crop yields. A $K_i > 1$ means crop i yielded more than expected on the basis of the ratio of the two crops, $\lambda_1 : \lambda_2$. Where $K > 1$ gives an overall yield advantage. Another measure commonly used is the *Aggressivity Coefficient (A)*. It measures competitive ability/dominance of one crop relative to the other when grown together and is defined by

$$A_{1/2} = \frac{\bar{y}_{i(j)l_i..m}}{\bar{y}_{i.s}\lambda_1} - \frac{\bar{y}_{j(i)l_j..m}}{\bar{y}_{j.s}\lambda_2} = \left(\frac{L_i}{\lambda_1}\right) - \left(\frac{L_j}{\lambda_2}\right)$$

Positive values would indicate that crop i is more competitive than crop j in an intercrop. Indices of competition are important in determining the compatibility of crops for intercropping.

Economic returns (RE)

Economic returns are directly related to total yield of an intercrop. They are computed as $p_1\bar{y}_{i(j)l_i..m} + p_2\bar{y}_{j(i)l_j..m}$ or

$$RE = \bar{y}_{i(j)l_i..m} + R\bar{y}_{j(i)l_j..m} \tag{2.7}$$

where $\bar{y}_{i(j)l_i..m}$ and $\bar{y}_{j(i)l_j..m}$ are mean yields as defined previously and p_1 and p_2 are the market prices of crop *1* and *2*, respectively. The R value is the ratio of p_2/p_1. The economic return from each intercrop system is computed and compared with that of sole cropping. The use of economic returns should also incorporate costs of inputs to give an indication of profits. Other measures of intercropping benefits that can be evaluated like economic returns are protein value, calorie value, farmers' value of the two crops etc., and in these cases, R can be a ratio of protein or calorie or farmers' values, respectively.

All these indices or measures can be applied to the same intercropping experiment to answer particular questions. For example, for land use efficiency, LER and ATER would be appropriate; for profit optimization, RE would be the most appropriate; and indices of competition are important in determining the compatibility of growing two or more crops or different crop density levels together. The weight attached to any of these indices depends on researcher's objectives. In the present study, interest is centred on LER, yield and RE since the discussion based on these measures can be generalised to the others. It is to be noted that many other responses apart from yield are also used to evaluate the cropping systems. These include among others soil erosion control, soil structure improvement, insect and disease control, nitrogen fixation, etc.

The effective evaluation of these measures requires appropriate experimental designs to be used. The next section is a review of some of the common designs that have been employed to enable evaluation of the above mentioned advantages of intercropping.

2.1.4 Some experimental designs common in intercropping

It has been noted that less attention is put on design and statistical analysis for data involving more than one crop, i.e. in intercropping experiments (Mead and Riley, 1981; Federer, 1993). The main purpose of this section therefore is to verify this statement by a review of experimental designs commonly used in intercropping studies reported in literature. Two surveys were conducted; one survey was conducted in Uganda in 2001 and another was conducted on literature reports in the *Journal of Agricultural Science Cambridge* and the *Journal of Experimental Agriculture.*

Most of the intercropping work done at International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), UK, for instance, used simple designs (Mead and Riley, 1981). Over 90% of the intercropping studies reported in the *Journal of Experimental Agriculture* for the last 20 years had a simple treatment structure with one or two factors (Mead and Riley, 1981). Apart from the few systematic designs and two cross-criss designs, the rest used either a simple randomised block or a split-plot design (Mead and Riley, 1981).

The results from the survey on intercropping experiments reported in the *Journal of Agricultural Science, Cambridge* and the *Journal of Experimental Agriculture* for the last 30 years are reported in Table 2.1. A total of twenty eight articles and thirty three articles in the *Journal of Agricultural Science, Cambridge* and the *Journal of Experimental Agriculture*, respectively were concerned with intercropping studies. In these articles more than 60% used RCBD or split-plot design. Less than 10% of them used other designs such as bivariate factorial design and systematic design. A survey was also conducted in Uganda on experimental designs used in intercropping experiments carried out by National research institutes and Makerere University over the period of ten years (1990 - 2001). The survey covered two national research institutes that are involved in annual and biennial crop research. Review of technical reports and theses

was done. Personal interviews with the agronomy researchers and graduate students of agronomy was also done. Of the 59 experiments reported, 53 were carried out using ordinary RCBD or split-plot in RCBD and 5 used replacement series and none used systematic design, as shown in Table 2.1.

Table 2.1: Survey results on types of designs used in intercrops

Design type	Frequency in a survey		Percentages	
	Uganda	Journals	Uganda	Journals
RCBD	38	33	64.41	54.10
RCBD with replacement series	5	4	8.47	6.56
RCBD with split plots	16	19	27.12	31.15
Other designs	-	5	-	8.19
Total	59	61	100.00	100.00

Many existing experiments on intercropping, which include two or more factors, use split-plot designs. At the International Institute of Tropical Agriculture (IITA), the most commonly used designs for intercropping experiments are the ordinary RCBD and the split plot designs (cited from Federer, 1993). Other designs that have been found very useful for intercropping studies are the systematic designs (Federer, 1993; Mead and Riley, 1981). However, these designs have not been widely used (Mead and Riley, 1981). Mead (1994) comments that these designs are an extremely useful addition to the practical statisticians library of designs and they should not be discarded because they do not satisfy the general principles that treatment allocation should be randomised, or that the number of treatment levels should not exceed four. Mead (1994) explains that many statisticians who are aware of the importance of randomisation in the analysis of experimental data, have been slow to accept the advantages of the

use of systematically arranged factor levels. He believes that this is based on narrow view of experimentation and of methods of analysis which are appropriate for drawing conclusions from experimental data.

2.2 Modelling responses

2.2.1 Nature of responses

Most of the analyses associated with measures of intercropping advantages that appear in literature are deterministic. In this section, therefore, the main focus is to consider some statistical models for density-yield relations to model these measures. The developments in this section are based on a RCBD. Consider a simple experiment consisting of 3 factors, namely a densities of crop i, p spatial arrangements and b densities of crop j. Assume r replicates. The responses that are possibly of interest for evaluating different treatment combinations can be categorised into two, namely observed and derived. The observed responses include yields of crop i and crop j and their respective yield components, disease incidence, pest infestation, etc. The derived responses include total yield, partial LER's, total LER, relative LER, economic returns, calorie values, area-time equivalent ratio (ATER), etc.

From the range of equations in Section 2.1.3, the derived responses are expressed as linear combinations of yield. For example, consider the following two general alternative equations,

$$y = \alpha_1 y_{i(j)l_i l_j h} + \alpha_2 y_{j(i)l_i l_j h} \tag{2.8}$$

$$y = y_{i(j)l_i l_j h} + \lambda y_{j(i)l_i l_j h} \tag{2.9}$$

where $y_{i(j)l_i l_j h}$ and $y_{j(i)l_i l_j h}$ are the yields of crop i and crop j respectively, and α_1, α_2 and $\lambda = \frac{\alpha_1}{\alpha_2}$ are coefficients of interest. For instance, y is the LER value for the case of

(2.8) and relative LER value for the case of (2.9) when the coefficients α_1 and α_2 are reciprocals of sole crop yields. If these coefficients were the prices of the crops, then y is the economic returns value, and y will be calorie value if the coefficients were the caloric values of the two crops. These coefficients can also be farmers' values of the two crops, coefficients from multivariate analysis such as bivariate analysis of component yields, etc. If the yield data of the two crops are normally distributed, it follows that their linear combination are also approximately normally distributed.

2.2.2 Modelling of the responses

The above mentioned responses can be categorised as follows;

- Univariate - Single crop analysis

  (1) Yield and each yield component taken singly (2) Partial LERs

- Univariate - combined crop analysis

  (1) Total yield/combined yield (2) Total LER, Relative LER, RE, ATER, etc.,

- Multivariate

  (1) Bivariate yield analysis (2) Yield components of each crop

The mathematical modelling of these responses as a function of plant populations may take various forms. Some of the suggested models given below are based on model (2.3).

Modelling yields

Various models can be employed to model intercropping yields or yield components. Some of these models are now discussed below.

Model I: *Regression or Response curve model*

The following regression model can be fitted at each hth level of spatial arrangement.

$$y_{hi(j)kl_il_j} = \beta_{0ik} + \beta_{1ik}d_{il_ik} + \beta_{2ik}d_{jl_jk} + \beta_{11(j)k}(d_{il_ik}.d_{jl_jk}) + \varepsilon_{hi(j)l_il_jk} \tag{2.10}$$

where $i = 0, 1, 2, \ldots, a, \quad j = 0, 1, 2, \ldots, b, \quad h = 1, 2, \ldots, r$ and $k = 1, 2, \ldots, p$. Note that $l_j = 0$ means a pure stand of crop i and the combination $(l_i = 0, l_j = 0)$ does not exist, where $y_{hi(j)kl_il_j}$ is the yield for l_ith density level of crop i intercropped with density level l_j of crop j in kth spatial arrangement. Also d_{il_ik}, d_{jl_jk} and $d_{il_ik}.d_{jl_jk}$ are levels of densities of crop i, crop j and their interactions, respectively. A quadratic effect in d_{il_ik} can be included. This model (2.10) treats the two crop densities as quantitative factors and spatial arrangements as a qualitative factor. The unknown regression parameters β_{1ik}, β_{2ik} and $\beta_{11(j)k}$ are concerned with the effect of crop l_i density levels, l_j density levels and their interactions (additive effect). This regression model could also be expressed using orthogonal polynomials since plant densities can be taken as quantitative factors. Model (2.10) is an ordinary linear regression model and thus least squares or maximum likelihood methods can be used to estimate the unknown parameters β_{0ik}, β_{1ik}, β_{2ik} and $\beta_{11(j)k}$. Graphical presentation of the response surfaces/curves facilitates the identification of the maximum of the response.

Model II: *Factorial linear model*

The data can be analysed as a 3-way factorial design on the yield (or any other response variable) for each crop. For example, the yield of crop i can be analysed as an $(a-1) \times b \times p$ factorial. Only $(a-1)$ levels of crop i are involved since the monocultures of crop j would not contain yield responses of crop i. The factorial linear model can be

expressed in the form

$$y_{hi(j)k} = \mu + \alpha_k + \tau_i + \gamma_j + \beta_h + (\tau\alpha)_{ik} + (\tau\gamma)_{ij} + (\gamma\alpha)_{jk} + (\alpha\gamma\tau)_{ijk} + \varepsilon_{hi(j)k} \quad (2.11)$$

where β_h is the hth replicate effect, α_k is the kth spatial arrangement effect, τ_i is the effect of density levels of crop i and γ_j is the effect of the density levels of crop j, where $(\tau\alpha)_{ik}$ is the interaction of spatial arrangement and density levels of crop i, $(\gamma\alpha)_{jk}$ is the interaction of spatial arrangement and density levels of crop j and $(\tau\gamma)_{ij}$ is the interaction of densities of crop i and crop j. The remaining term $(\alpha\gamma\tau)_{ijk}$ is a three-way interaction term. This model can also be extended to factorial orthogonal polynomial models where a linear regression effect, quadratic linear effect, etc. for each crop density factor and their interactions are computed.

Model III *Regression model with additive effect term*

Yield or any yield component of each crop can be analyzed separately as a function of density using the model

$$y_{i(j)l_il_jhk} = \beta_{0ik} + \beta_{1ik}d_{il_i} + \gamma_{i(j)k}(d_{il_i}, d_{jl_j}) + \varepsilon_{i(j)l_il_jhk} \quad (2.12)$$

where $y_{i(j)l_il_jhk}$ is the yield of crop i grown at density l_i in hth replicate together with density level l_j of crop j in spatial arrangement k. The unknown parameters β_{0ik} and β_{1ik} are the intercept and the slope, respectively. The term $\gamma_{i(j)k}$ is the measure of intercrop additive effect of crop j on crop i in kth spatial arrangement. The assumption in this set up is that yield observations are normally distributed and $\varepsilon_{i(j)l_il_jhk} \sim \text{iid}(0, \sigma^2)$. By applying generalised least squares theory (Federer, 1993), best linear unbiased estimates of the following parameters are obtained as

$$\tilde{\beta}_{0ik} = \bar{y}_{11..} - \bar{d}_{i.}\tilde{\beta}_{1ik} \quad (2.13)$$

$$\tilde{\beta}_{1ik} = \sum_{i=1}^{a}\left(d_{il_i} - \bar{d}_{i.}\right)\left(\bar{y}_{11l_i.} - \bar{y}_{11..}\right) / \sum_{i=1}^{a}\left(d_{il_i} - \bar{d}_{i.}\right)^2 \quad (2.14)$$

$$\tilde{\gamma}_{i(j)}(d_{il_i}, d_{jl_j}) = \bar{y}_{i(j)d_{il_i}d_{jl_j}} - d_il_i\tilde{\beta}_{1ik} - \tilde{\beta}_{0ik} \quad (2.15)$$

where $\bar{y}_{11..}$ is sole crop i mean yield. The parameters $\tilde{\gamma}_{1(2)}(d_{1l_i}, d_{2l_j})$, $\bar{y}_{1(2)d_{1l_i}d_{2l_j}}$, $\tilde{\beta}_{11k}$ and $\tilde{\beta}_{01k}$ are obtained for each spatial arrangement. The parameters from different spatial arrangements can be compared since each parameter estimate has its variance or can be subjected to two-way analysis of variance with replicate and spatial arrangements as two classifications. The t-test can also be conducted on additive effects. This idea will be explored further in Chapter 4. A plot of $\tilde{\gamma}_{i(j)}(d_{il_i}, d_{jl_j})$ (also represented as $\tilde{\gamma}_{i(j)}(l_i, l_j)$) parameter estimated against plant densities could be done and the pattern observed. The variances of additive effects can be computed using the following equation (Federer, 1993)

$$\mathrm{Var}\big(\tilde{\gamma}_{i(j)}(l_1, l_2)\big) = \frac{\sigma_i^2}{r}\left(1 + \frac{(d_i l_i - \bar{d}_{i.})^2}{\sum_{li}(d_i l_i - \bar{d}_{i.})^2} + \frac{1}{n_i}\right) \tag{2.16}$$

$$\mathrm{Cov}\big(\tilde{\gamma}_{i(j)}(l_1, l_2), \tilde{\gamma}_{i(j)}(l_1', l_2')\big) = \left\{\sum^{n_i} \frac{d_{il_i}}{n1} + d_{il_i}d_{il_i'} - (d_{il_i} + d_{il_i'}d_{i.})\right\}\left\{\frac{\sigma_1^2}{r\sum^{ni}(d_{il_i} - \bar{d}_{i.})^2}\right\} \tag{2.17}$$

where d_{il_i} is the l_ith density level of crop i, n_i is the number of density levels used to compute intercrop yield means, and r is the number of replications.

This model is illustrated on the following example.

Example 2.2.1 The data used in this example are from an intercropping experiment on population studies of four densities of simsim ($s1 - s4$) and four densities of finger millet ($f1 - f4$). The experiment was laid out in split plots in RCBD and replicated three times. The main interest is to determine which density level combination gives the maximum yield response. Included in Table 2.2 are the mean yields of simsim. This experiment was carried out in Uganda by a student at Makerere University. The data is included in Appendix D. The presented density levels are plant densities per hectare.

Consider fitting a simple regression model to simsim yields in pure stand

$$y_{11ih} = \beta_{01} + \beta_{11}d_{ih} + \varepsilon_{11ih} \tag{2.18}$$

Table 2.2: Mean yields (Kg/ha) of simsim from 3 replicates

	Simsim densities				
F. millet densities	*44,400* (s1)	*22,200* (s2)	*11,100* (s3)	*6,330* (s4)	TOTAL
0 (f0)	723.23	488.90	450.00	352.23	2,014.36
10,000 (f1)	305.57	260.00	250.00	215.57	1,031.13
20,000 (f2)	493.33	387.40	316.03	260.00	1,456.76
40,000 (f3)	282.97	238.50	200.00	162.20	883.667
60,000 (f4)	419.97	333.33	253.33	213.33	1219.97
Total	2,225.06	1,708.13	1,469.36	1,203.33	6,605.89

from which the estimates $\hat{\beta}_{01} = 503.59$ and $\hat{\beta}_{11} = 0.00974$ are obtained. Also $\bar{d}1. = 21008$. The fitted mean yields for sole crop are computed using these estimates. These are

$$\begin{aligned}
\overline{y}'_{11s1} &= 503.59 + 0.00974(44400 - 21008) &= 731.43 \\
\overline{y}'_{11s2} &= 503.59 + 0.00974(22200 - 21008) &= 515.20 \\
\overline{y}'_{11s3} &= 503.59 + 0.00974(11100 - 21008) &= 407.09 \\
\overline{y}'_{11s4} &= 503.59 + 0.00974(6330 - 21008) &= 360.63
\end{aligned}$$

Using expression 2.15, the values of $\gamma_{1(2)(l_1,l_2)}$ ($l_1 = 6330$, 11100, 22200, 44400 and $l_2 = 0$, 10000, 20000, 40000, 60000) are computed. Some of these estimates are as follows

$$\begin{aligned}
\hat{\gamma}_{1(2)}(44400,\ 0) &= 723.23 - 731.43 = -8.20 \\
\hat{\gamma}_{1(2)}(44400,\ 10000) &= 305.57 - 731.43 = -425.86 \\
\hat{\gamma}_{1(2)}(22200,\ 0) &= 488.90 - 515.20 = -26.30 \\
\hat{\gamma}_{1(2)}(6330,\ 10000) &= 215.57 - 360.63 = -145.07
\end{aligned}$$

The summarised results of the additive effects are provided in Table 2.3.

Table 2.3: Additive effects ($\gamma_{1(2)(l_1,l_2)}$)

	$\hat{\gamma}_{1(2)(44400,l_2)}$	$\hat{\gamma}_{1(2)(22200,l_2)}$	$\hat{\gamma}_{1(2)(11100,l_2)}$	$\hat{\gamma}_{1(2)(6330,l_2)}$
$l_2 = 0$	-8.20	-26.30	42.91	-8.41
10,000	-425.86	-255.20	-157.09	-145.07
20,000	-238.10	-127.82	-133.97	**-100.63**
40,000	-448.50	-276.71	-207.09	-198.43
60,000	-311.50	-181.87	-153.76	-147.30

Variances of these estimates can be computed and comparisons can be easily made. Results from Table 2.3 suggest that that the pure stand of simsim performs better than intercrops. Also the intercrop combination (6330, 20000) is the best among all intercrop combinations since the yield of sole simsim is reduced the least in this combination. Consider comparing the maximum additive effect of 60,000 finger millet plants on simsim (i.e. -147.30) with the best intercrop combination (i.e. -100.63). Their variances and covariance are computed using (2.16) and (2.17) where $\text{Var}(\hat{\gamma}_{1(2)(6330,\,20000)}) = \text{Var}(\hat{\gamma}_{1(2)(6330,\,60000)})$ given by

$$\frac{8120}{3}\left(1+\frac{(6330-21008)^2}{(6330-21008)^2+\ldots+(44400-21008)^2}+\frac{1}{4}\right)=0.5\times 8120$$

and their covariance computed similarly is 0.022×8120. From these results a calculated t value (to test the whether there is significant difference in these additive effects) is 0.5237, i.e. from

$$t=\frac{\hat{\gamma}_{1(2)(6330,\,60000)}-\hat{\gamma}_{1(2)(6330,\,20000)}}{\sqrt{\text{Var}\left(\hat{\gamma}_{1(2)(6330,\,60000)}-\hat{\gamma}_{1(2)(6330,\,20000)}\right)}}=\frac{(-100.63--147.30)}{\sqrt{(0.5\times 8120\times 2+0.022\times 8120)}}=0.52$$

Then using the t-test with 38 degrees of freedom ($(r-1)n_i+(r-1)m_1m_2$ with m_i's are intercropped density levels for crop i), the calculated value is not significant at $\alpha = 0.05$. Therefore the two additive effects are not significantly different. Analysis of variance

procedure in GenStat was also conducted on this data and the These conclusions agree with those arrived at using ANOVA and treating plant densities as factors (where LSD of 166.3 was used to test the yield mean difference corresponding to these effects and concluded, not significant; Table I - Appendix A). A quadratic effect could also be fitted instead of only a simple linear function in d_i in (2.18)

Model IV: *Bivariate Analysis*:

Analysis can be done on similar observed variables on both crops jointly. Here crops are used as variates and sole crop responses do not enter the analysis. The commonly used response vectors are the yields of the two crops. This is referred to as bivariate analysis of crop yields. In the present case the new response vector can be given as $\mathbf{y} = \left(y_{hi(j)k},\, y_{h(i)jk}\right)$ where $y_{hi(j)k}$ is the yield of crop i and $y_{h(i)jk}$ is the yield of crop j. A linear combination (canonical variable or discriminant function) of the two yields are formed from MANOVA (multivariate analysis of variance). Univariate ANOVA can then be conducted on this variable/function or a graphical presentation of results as given by Pearce and Gullivar (1979) can be applied. Further details about this method can be found in Pearce and Gullivar (1979) and Mead and Riley (1981). In the present study, since the main interest is to study the response pattern and since this method emphases significance testing, this method will not be pursued any further.

Modelling LER

Federer (1993) suggested analysis on relative LER's, since the distribution of relative LER approximates more closely the distribution of the yield than that of ordinary LERs. Assuming a normal bivariate distribution for the yields of the two crops and taking LERs as random parameters or linear combination of the two yields, modelling LER is possible. Partial LERs (L_A) for crop i can be taken as a function of density and thus replaces $y_{hi(j)k}$ in models (2.10) or (2.11).

Total LER or relative LER can be modelled as a function of *total* plant population (T). Thus,

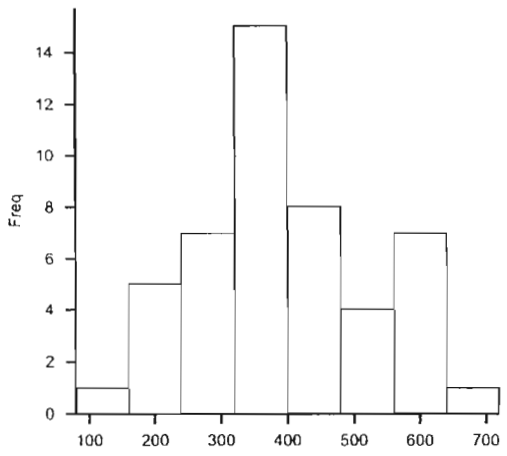
$$L_{hij} = \mu + \alpha T_i + \gamma S_j + \lambda(TS)_{ij} + \varepsilon_{hij} \qquad i = 1, 2, \ldots, n \quad j = 1, 2, \ldots, p \qquad (2.19)$$

where L_{ij} is the relative LER for the ith total population (T) from the jth spatial arrangement (S). In the case where crop i has a constant crop density level, $a = 1$, the effect of changing densities of crop j on LER can be modelled as a function of these densities. Thus, for this case relative LER replaces $y_{hi(j)kl_il_j}$ in (2.12). This model (2.19) is only limited to intercrop responses as it does not provide for sole crop responses.

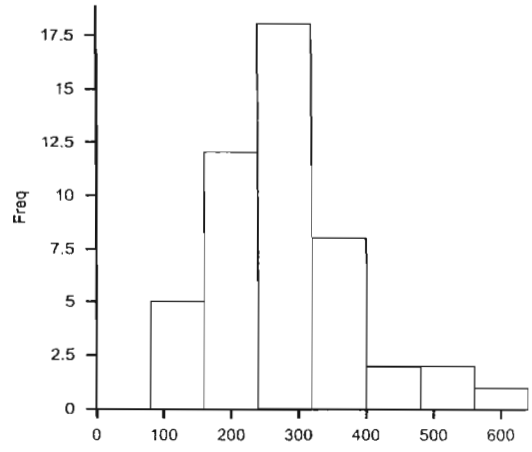
Consider for example the data in Example 2.2.1. Using the averages of pure stand yields as standardising factors, the relative LER values can be computed. The relative LER's have a similar distribution to that of the two crop yields (Figure 2.1). Figure 2.1(c) demonstrates that the distribution of relative LER is much closer to that of the yields (Figure 2.1 (a) and (b)). The advantage of modelling relative LER is that both crop yields are used for efficient intercrop evaluation.

Modelling economic returns (RE)

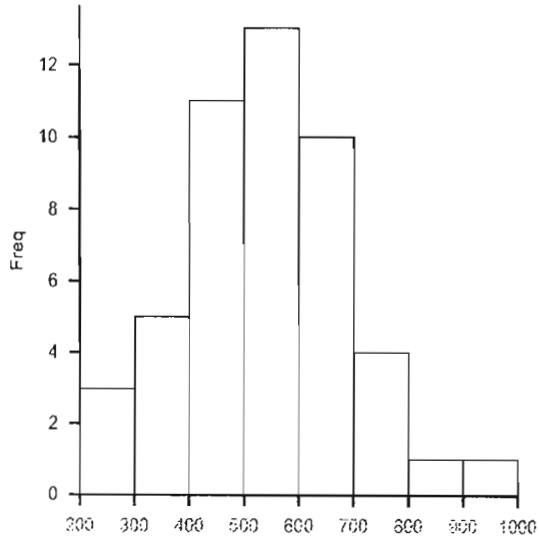
The economic returns can be modelled using the above models by simply replacing $y_{hi(j)kl_il_j}$ in (2.10) or (2.11) or by replacing L_{hij} in (2.19) with RE. The assumptions made above and the limitations cited also apply for RE model. Assuming ratio of the cost of simsim to finger millet is 1.5:1, RE values were computed. Their distribution like relative LER is the similar as that of yields (Figure 2.1(d)). Although the histogram for simsim yields suggests that simsim data do not have an exact normal distribution, the idea of presenting the histograms is to illustrate how the distribution of LER and RE approximates closely to that of the crop yields.



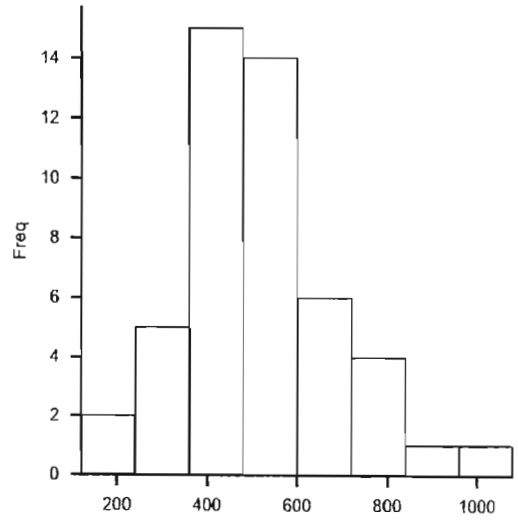
(a) Finger millet yield



(b) Simsim yield



(c) Relative LER



(d) RE

Figure 2.1: Data histograms of simsim and finger millet yields, LER and RE

General comments

The general matrix form of models I, II and III is given as

$$\mathbf{y} = X\boldsymbol{\beta} + \mathbf{e} \tag{2.20}$$

where $\mathbf{y}$ is a vector of observations $\boldsymbol{\beta}$ is the vector of parameters corresponding to plant populations and blocks X is the design matrix and $\mathbf{e}$ is an error vector. In the model-based analysis block effects can be considered as random effects and this leads to a model

$$\mathbf{y} = X\boldsymbol{\beta} + Z\mathbf{u} + \mathbf{e} \tag{2.21}$$

where $\boldsymbol{\beta}$ is the vector corresponding to fixed effects (factors consists of all levels of interest e.g. plant populations levels) and X is a design matrix, $\mathbf{u}$ is a vector of coefficients corresponding to random effects (factors or factor levels selected at random from a population e.g. main plot or block effects) with a design matrix Z. This is referred to as a linear mixed model (Henderson, 1990). Most of the discussions in this study are based on (2.21).

In this chapter the feasibility of handling intercropping data using the general linear model has been demonstrated. A clear understanding of the response variable from intercropping experiments is required in deciding on an appropriate model for analysis. More importantly, it has been demonstrated that LER and RE in plant population studies can be analysed in the same way as crop yield.

Chapter 3

Randomised and Systematic Designs

3.1 Background

The primary objective of most agricultural field experiments is the unbiased and efficient estimation of treatment effects, contrasts, and efficient exploration of the response pattern. A broad inference space based on the results from such experiments is always sought. In other words the conclusions arrived at from an experiment in one agricultural field should apply to apply to other fields as well (Yates, 1939). Three key concepts, namely replication, blocking and randomisation are fundamental in experimental designs to enable this inference space. Contrary to the concept of randomisation is the systematic allocation concept in experimental designs. The role of both randomisation and systematic allocation in intercropping experiments in particular is addressed in the subsequent sections.

Replication involves application of the same treatment independently to several alike experimental units (plots) under identical conditions. Replication is necessary for the

estimation of experimental error. Furthermore, it increases precision of treatment estimates. The experimental material, divided into units to which a number of treatments are to be applied are rarely sufficiently homogeneous. By appropriate blocking, each block will contain uniform units. By allocating each treatment within uniform blocks, units can be assumed to have constant mean within the blocks. Blocking controls variation due to local heterogeneity hence reducing the experimental error. Randomisation is the process of randomly allocating or assigning treatments to experimental units. The process is intended to average out systematic effects present in the field due to extraneous factors that are not under the control of the experimenter at the time of experiment layout. The concept of randomisation is a single most crucial principle in the design of experiments and hence merits further discussion.

3.2 Randomisation theory

Randomisation was one of the ideas R.A. Fisher introduced in experimental design and analysis in the 1920's and 1930's. The main reason for randomisation is to neutralise spatial or temporal dependence that occurs in field observations. This ensures valid statistical analysis (i.e. in those methods of analysis where spatial correlation is not modelled). Prior to the work of Fisher, treatments have been assigned to experimental units either on a systematic or on a subjective basis, without considering in consequence the heterogeneity of the field. Such allocations of treatments may lead to systematic error in the estimation of treatment effects and error variance and may lead to bias in the results. Bias is introduced when one particular treatment is systematically assigned to a better environmental conditions than the others. The comparison of such a treatment with others does not only reflect the treatment differences but also the environmental effect. This effect is minimised when treatments are randomly assigned to the experimental units and any systematic effects present in the field tend to aver-

age out. Randomisation tends to eliminate the influence of extraneous factors (such as environmental factors) which are not under the direct control of the experimenter and also precludes the systematic error in estimation of error variance and treatment effects.

Randomisation neutralises spatial dependence by imposing a uniform/constant correlation structure over all the possible permutations of the treatments on the designs. As an example, the randomisation theory of complete block design is illustrated.

Consider n experimental units. Since each unit/plot (i) can receive only one treatment, the yield of treatment k from block j is

$$y_{jk} = \mu + \beta_j + \tau_k + \sum_{i=1}^{t} \delta_{ij}^k \epsilon_{ij} \qquad j = 1, \ldots, r, \;\; k = 1, \ldots, t \tag{3.1}$$

where

$$\delta_{ij}^k = \begin{cases} 1 & \text{if unit } i \text{ in block } j \text{ receives treatment } k \\ 0 & \text{Otherwise} \end{cases}$$

μ, β_j, and τ_i denotes overall mean, block effects and treatment effects, respectively.

When treatments are assigned at random within each block, the δ_{ij}^k are random variables whose joint probability distribution is induced by the randomisation (Kempthorne, 1952). They are characterised by

1. $E[\delta_{ij}^k] = \frac{1}{t}$

2.

$$E[\delta_{ij}^k \delta_{i'j'}^{k'}] = \begin{cases} t^{-1} & \text{for } (i,j,k) = (i',j',k'). \\ t^{-2} & \text{for } i \neq i' \text{ and any } j,\; j' \\ (t(t-1))^{-1} & \text{for } j = j',\; k \neq k', i \neq i' \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

If the effect of experimental units are taken independently of the treatments, as $\omega_{jk} = \sum_{i=1}^{t} \delta_{ij}^k \epsilon_{ij}$, Grondona and Cressie (1991) outlined the following points

1. $\mathrm{E}(\omega_{jk}) = E_W E_R\big(\omega_{jk}|\{W_{ij}; i = 1, \ldots, t\}\big) = 0$

2.

$$Cov[\omega_{jk}\,\omega_{j'k'}] = \begin{cases} \sigma^2_{\omega_j} & j = j',\ k = k'. \\ -\sigma^2_{\omega_j}(t-1)^{-1} & j = j',\ k \neq k' \\ 0 & j \neq j' \end{cases} \tag{3.3}$$

where $E_w(\cdot)$ denotes expectation with respect to the probability distribution of the random process $W(\cdot)$ (smooth small-scale variation), and $E_R(\cdot)$ is the expectation with respect to randomisation distribution (Grondona and Cressie, 1991) where also

$$\sigma^2_{\omega_j} = \frac{1}{t}\sum_{i}^{t} E_W(\epsilon^2_{ij}) \tag{3.4}$$

From (3.3) and (3.4), it is clear that randomisation does not remove the correlation pattern but by making equally likely the fact that two treatments are adjacent, it neutralises it to a small negative uniform correlation between treatments within the same block. The null covariance between blocks can be understood intuitively since randomisation is performed independently for each block.

Cochran and Cox (1964) have compared randomisation to an insurance policy in that it is a precaution against disturbances that may or may not occur, and that may or may not be serious if they do occur. Randomisation enables the data analysis without the necessity of modelling the plot effects and makes valid the usual tests of significance. This is precisely the basis of design based inference.

A major drawback with randomised designs is that, when positive dependence exists, they are on average less efficient than good systematic designs (Martin, 1996). Watson (2000) notes that if spatial dependence is likely, the idea would be to use spatial design and spatial method of analysis. Williams (1952) suggested the use of systematic designs and data analysed by assuming errors to be correlated in a stationary linear

autoregressive process. If spatial dependence exists, the ordinary least square estimators will be unbiased but will be less efficient (Watson, 2000). In model-based analysis such as spatial analysis, the error structure is modelled directly and thus if the model is well specified it gives more efficient estimators i.e. generalised least square estimators. However, model-based analysis does not conflict with randomisation i.e. sound design but it is an adjunct to it. Nevertheless, unlike design-based inference, model-based inference is not bound by randomisation. It is upon this argument that model-based analysis, specifically spatial analysis of field experiments gained much attention. It is more plausible as an analytical tool because it is not always easy to know *a priori* what sort of variation will exist in the data.

Furthermore, in intercropping experiments the use of randomised block designs such as the ordinary RCBD for studies where many levels of plant densities are involved would result in blocks of enormous size. Experiments with large block sizes are not statistically desirable and are not very practical in terms of resources (Bleasdale, 1967). Very large sized blocks are associated with large error variances (Mead, 1994), especially in the tropics where variability is very high and would render a large experiment imprecise. The management of randomised designs in intercropping population studies is another factor for consideration. Systematic arrangement of levels of one crop easily facilitates management compared to randomised block designs.

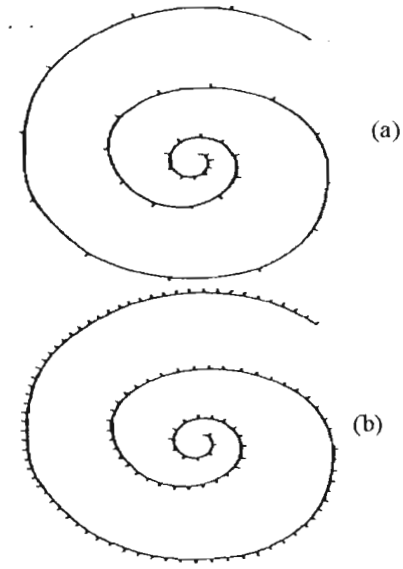
3.3 Systematic designs

3.3.1 General description

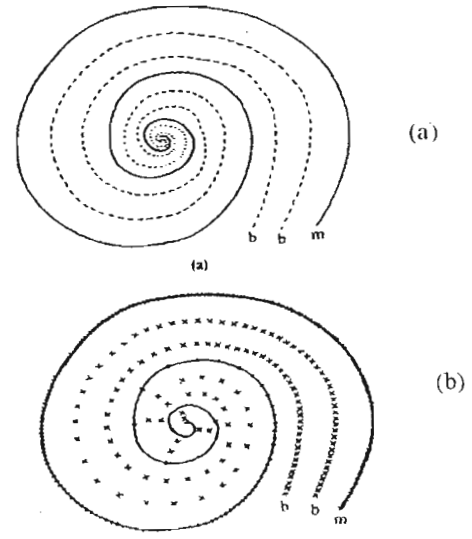
Systematic design in general refers to a design where allocation of treatments to experimental units is selected purposely, and is not an outcome of the valid randomisation scheme. In this study focus is given to systematic designs where quantitative factor

levels are arranged systematically, according to Nelder (1962). These designs are specifically for spacing experiments. This kind of systematic design consists of a grid of points, each representing the position of a plant, and having the property that the area per plant and/or the rectangularity of the space available to a plant changes in some consistent fashion over the different parts of the grid. Locally the design is assumed to be approximately rectangular, so that any plant has immediate neighbours in positions close to those that would occur in a strictly rectangular array. The grids are defined by straight lines or arcs of concentric circles, and the contours of equal area and equal rectangularity of arrangement are either straight lines or arcs of concentric circles (Nelder, 1962). Of the systematic designs suggested by Nelder (1962), the two that have been used in intercropping experiments include the 'fan' design (Huxley and Maingu, 1978; Federer, 1993) and parallel row arrangement (Willey and Rao, 1980). Many other forms of systematic designs apart from parallel row and fan design are also available. These include snail-shaped and circular-shaped systematic designs or Okigbo circular designs (Federer, 1993). In the Figures 3.1(I - IV) and 3.2(I - II), various forms of systematic designs are given. Different markings represent different crop species layout.

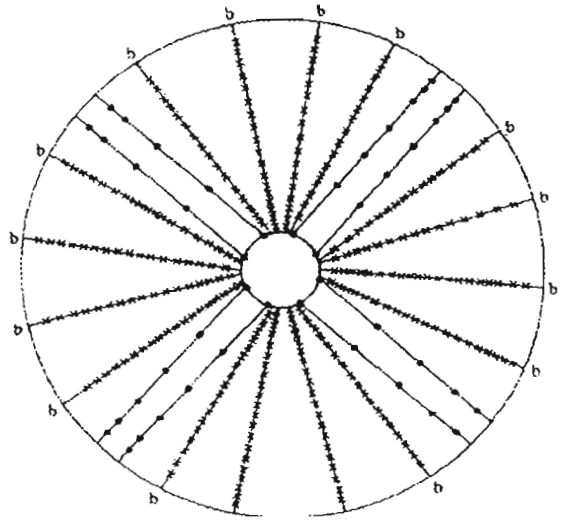
Systematic designs are used in experiments involving a spatial factor (spacing) and other factors (e.g. nutrients, genotypes, etc) that are applied to the main systematic plots (Mead, 1994). The systematic treatment of a spatial factor resembles that of a split-plot factor. The spatial factor is systematically arranged within 'subplots' whereas the other factors applied to 'whole plots' are randomised and mixed in the usual way (Mead and Stern, 1980).



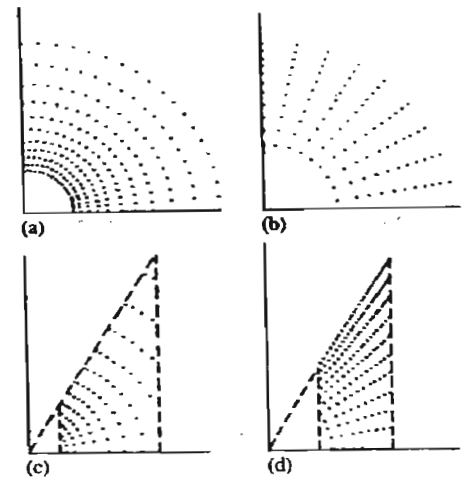
(I) Snail designs: Plant spacing within row (a) increasing, (b) constant



(II) Snail designs: (a) Density & row spacing varied, (b) density constant, row spacing varied.



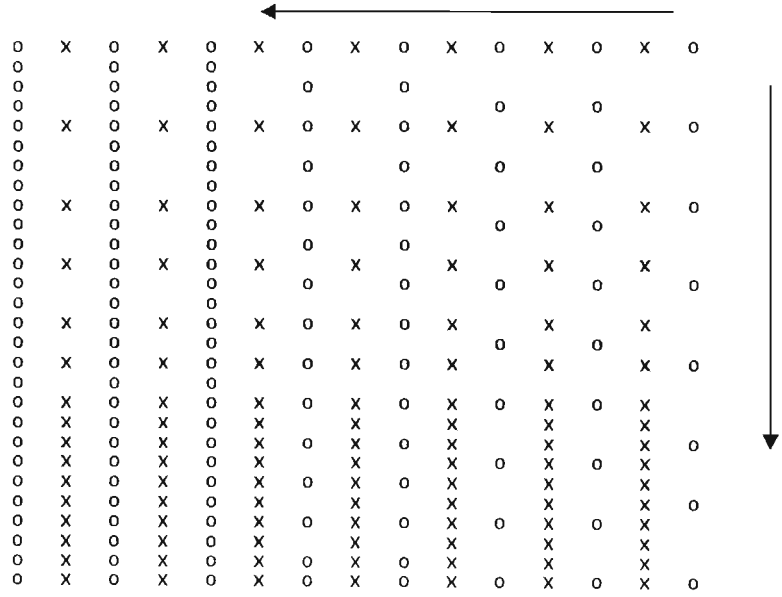
(III) Okigbo circular design; Two rows of one crop species interspersed with four rows of the second crop.



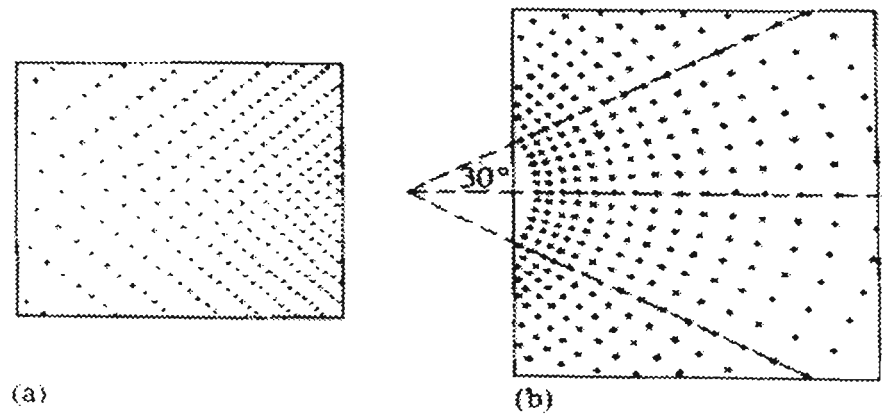
(IV) Nelder's fan designs: Four polar-coordinate grids. (a) & (b); shape fixed, (c) & (d); Equal-area contours

Source: Excerpts from Federer (1993, pp 214 - 219).

Figure 3.1: Snail, circular and fan systematic designs



(I) Two way systematic spacing for two crops (o and x) with densities varying perpendicularly (design I). (**Source**: Excerpt from Mead and Riley (1981))



(II) Rectangular fan-type designs. (a) Equal-area contours = vertical lines/horizontal lines
(b) Arrangement in a rectangular plot.
Source: Excerpt from Federer (1993, pp215)

Figure 3.2: Fan and parallel row systematic designs

3.3.2 Advantages of systematic designs

Recalling the arguments in Section 3.2, systematic designs have various advantages over randomised designs. These include efficient use of the available experimental material through reduction of planted and non-harvested areas. Secondly, systematic designs are not tied to a particular objective. This is important when prior information about parameters of a response model is little or not defined, a situation common in intercropping studies (Mead, 1979). It enables efficient response surface exploration as the inclusion of many plant population levels is possible on a small piece of land. In a systematic design, the levels of a quantitative factor vary systematically within the subplots, with a subsequent level being $10-15\%$ higher or lower than the preceding level. Each particular treatment level is surrounded by treatment levels differing only slightly, and therefore guard rows are unnecessary (Mead, 1994). Use of systematic designs give more information about the response surface than would do the conventional RCBD.

Consider, for example, Figure 3.3 obtained using the results in Table 2.3.

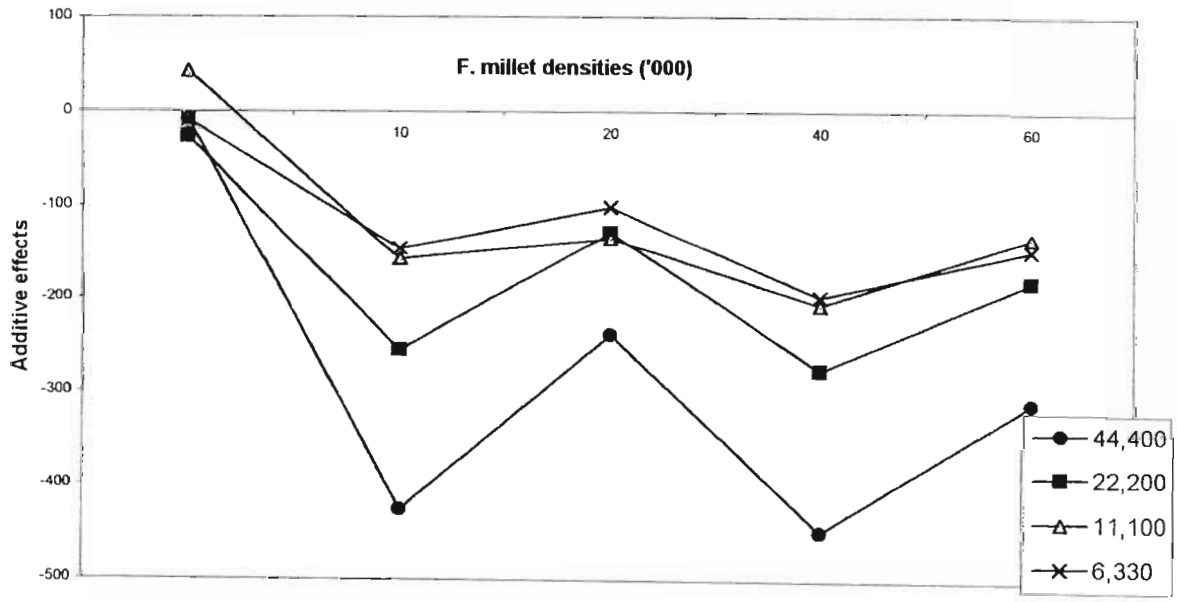


Figure 3.3: Plot of additive effects versus F. millet densities for each level of simsim densities

Suppose that due to laying the experiment in a RCBD, the finger millet density level of 60,000 plants per hectare was left out. Then the maximum for additive effect on simsim density of 11,100 plants per hectare would have been excluded. Also, the feature of the lowest additive effect for maize 44,400 plants per hectare would be missed if finger millet density level 40,000 was was not included but instead 60,000 plants per hectare included. The use of systematic designs which guarantee the span of all these density levels would also more likely guarantee spanning the maximum than RCBD. The counter argument is that if the density levels are selected carefully then the level that gives the maximum would be included. This argument may not hold most of the time because the range of densities in which the maximum lies in intercrops is not easily predictable from monoculture.

3.3.3 Disadvantages of systematic designs

A disadvantage associated with the use of systematic designs is the lack of randomisation. The assurance that estimates of 'error' may be unbiased, does not hold (Yates, 1939). Other disadvantages advanced by Yates (1939) are that the comparisons of different pairs of treatments are subject to different standard errors. Data from systematic designs are prone to spatial correlation and thus the conventional analysis of variance cannot always be used. The precision of treatment comparisons is greater for treatments falling on plots close together than those far apart in space (Steel and Torrie, 1990). In the case of intercropping studies the main interest is not in comparing subplot treatments but rather in the nature of the response pattern for which unbiased estimates of treatment effects and error variance can be obtained through modelling the spatial correlation. The error structure that may be associated with systematic designs is characterised by the correlation of spatial nature. Possible error structures include spherical and first order autoregressive structures.

Given a vector of correlated observations $\mathbf{y}$ and a model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ where $\boldsymbol{\beta}$ is a vector of treatment effects and the error term $\mathbf{e}$ with covariance matrix $V(\lambda)$ that depends on λ, the unknown spatial correlation parameters. According to Zimmerman and Harville (1989), and Bailey *et al.* (1995), if the distribution of the vector $\mathbf{y} - E(\mathbf{y})$ is symmetric, and if generalised least squares (GLS) is performed using the covariance matrix $V(\hat{\lambda})$ after $\hat{\lambda}$ is obtained using restricted maximum likelihood (REML) of a symmetric distribution, then estimates of treatment effects are unbiased. In line with this a simulation process was initiated to investigate the behaviour of the vector $\mathbf{y} - E(\mathbf{y})$ of data from systematic designs. A simulation of 400 Monte Carlo samples of 50 correlated random realizations assuming no treatment effects was conducted using PROC IML in SAS. The simulation code is included in Appendix B. These samples were analysed by modelling the correlation structure. A total of 89% of the samples produced means and medians that coincided. These results assure us that the estimates of fixed effects are unbiased. Furthermore, since the estimates of treatment effects are unbiased and also because the yield response from different main plots or blocks are independent, the analysis of variance is weakly valid (Bailey *et al.*, 1995).

Another disadvantage is that the pattern of the systematic arrangement coinciding with a field trend, for example with some field fertility pattern, may lead to biased estimates of error variance. Nonetheless, the proponents of systematic design argue that provided the blocking principle of design is maintained, any trend would affect a systematic design and randomised design to more less to the same degree (Mead and Riley, 1981). Cox (1951) indicated that for systematic designs, if treatment differences are completely orthogonal to the trend, then the estimation of error variance of treatment comparisons is estimated normally using least squares. If however, the treatment differences are not orthogonal to the trend, a valid estimate of the error variance can still be obtained although the error variance will be high i.e. there is loss in precision.

Consider, for example t treatments each replicated r times and suppose the fertility trend can be represented by a pth order polynomial, say

$$\alpha_1\xi_1' + \ldots + \alpha_p\xi_p' \qquad p < t$$

where ξ_j' is an orthogonal polynomial of order i for t equally spaced points and $\alpha_1, \alpha_2, \ldots, \alpha_p$ are coefficients. An example of this polynomial could be a four order polynomial in plots such as $0.2R_{(i)} + 0.6R_{(i)}^2 + o.12R_{(i)}^3 + 0.3R_{(i)}^4$ where $R_{(i)}$ = the plot position number in the row - average plot position number. Assuming the value of p is known, the population mean of the data measurement (y_k) on the ith plot is

$$a_k + \sum_{j=1}^{p} \alpha_j\xi_j' \quad k = 1, 2, \ldots, t$$

where a_k depends on only the kth treatment applied to the ith plot. Let $s_{jk}' \equiv \sum_k \xi_j'$, where $\sum_k$ means the trend is summed over all the plots receiving treatment k. The treatment differences are completely orthogonal to the trend if $s_{jk}' = 0$ and $s_{jk}' \neq 0$ if there is no orthogonality. Cox (1951) has shown that in the presence of non-orthogonality, to estimate a_k, *normalised* orthogonal polynomials ξ_j that are simple multiples of ξ_j' are introduced such that

$$\sum \xi_j = 0, \ \sum \xi_j\xi_{j'} = \delta_{jj'} = \begin{cases} 1 & j = j' \\ 0 & j \neq j' \end{cases} \tag{3.5}$$

where summation is over all the plots. The trend can then be represented as $\sum_{j=1}^{p} \alpha_j\xi_j$, and the sum of squares to be minimized is then

$$\sum_{k=1}^{t} \sum_{k} \left(y - a_k - \sum_{j=1}^{p} \alpha_j\xi_j \right)^2$$

The least squares equations according to constraints in (3.5) are

$$-2\sum_{k=1}^{t} \sum_{k} \left(y - a_k - \sum_{j=1}^{p} \alpha_j\xi_j \right) \tag{3.6}$$

$$-2\sum_{k=1}^{t} \sum_{k} \left(\xi_j y - \xi_j a_k - \sum_{j} \alpha_j\xi_j\xi_{j'} \right) \tag{3.7}$$

Cox (1951) showed that these equations can be expressed alternatively as $\mathbf{Kb} = \mathbf{y}^*$, where $\mathbf{K}$ is a $(t+p)$th order square matrix

$$\begin{pmatrix} r\mathbf{I}_t & \mathbf{S} \\ \mathbf{S} & \mathbf{I}_p \end{pmatrix},$$

$\mathbf{y}^*$ is a column vector of $\{\sum_1 y, \ldots, \sum_t y, \sum \xi_1 y, \ldots, \sum \xi_p y\}$ while $\mathbf{b}$ is a vector of $\{\hat{a}_1, \ldots, \hat{a}_t, \hat{\alpha}_1, \ldots, \hat{\alpha}_p\}$. Matrix $\mathbf{I}_t$ is a tth order unit matrix, $\mathbf{S}$ is a $t \times p$ matrix whose (i,j)th element is $\sum_k \xi_j$. Solving these equations for $t = 2$ for example, yields

$$\text{var}(\hat{a}_1 - \hat{a}_2) = \frac{2\sigma^2}{n}\left\{1 - \frac{2}{n}\sum_{j=1}^{p} s_j^2\right\}^{-1} \tag{3.8}$$

From (3.8) it is clear that a valid estimate of variance for comparing treatment effects is obtainable though there is loss in the efficiency of comparing the treatments by a factor of $\{1 - \frac{2}{n}\sum_{j=1}^{p} s_j^2\}$. Again, the argument here is that in practice the designs will be near orthogonality.

In intercropping experiments involving plant populations, systematic designs can be used and the trend, error or both can be modelled to recover information which would otherwise be lost because of systematic arrangement of plant densities. Through modelling of spatial correlation and/or trend, less biased estimates and valid estimates of treatment effects and error variance are obtainable. Secondly, since the main interest in these experiments is in the understanding of the response pattern and not distinguishing individual plant population levels, this can be achieved more efficiently using systematic designs. In the section that follows, the scenarios in which systematic designs can be used are enumerated.

3.3.4 Some scenarios for use of systematic designs

Three scenarios in which systematic designs can be used in intercropping experiments are described. These scenarios are denoted as design I, II and III for ease of reference

in the subsequent sections. The descriptions are set up parallel to ordinary split plot in RCBD for elaboration purposes. Scenario I or design I applies when both crops to be grown are main crops and both have varying densities. Design II is also applicable when the crop densities of the two crops are varying. Design III is applicable when one crop density is fixed and the other crop density is varying.

Scenario I:

In a split-plot arranged in a RCBD, different density combinations are allocated randomly to subplots while spatial arrangements are randomly made to the whole plots. In a systematic design, the same procedure is followed except that the levels of crop densities are arranged systematically in systematic 'subplots'; one crop density arranged systematically column-wise while the other crop density levels arranged systematically row-wise i.e. perpendicular to each other. The direction of systematic arrangement being chosen at random in each 'main plot'. This design will be referred to as design I (Figure 3.2(i)).

This setting is most applicable when both crops are of main interest. For example, planting 8 densities of beans to columns and 6 densities of maize to rows. The most probable analyses could be by evaluating the effect of one crop on another and determining optimum yield or partial LER and also jointly modelling relative LER or RE. Among many others, model (2.12) can be applied for this case. The assumption is that the yield observations are distributed normally.

Scenario II:

In a split-plot in a RCBD, different density levels of crop j are allocated randomly to subplots while a combination of crop i densities and spatial arrangements, and other factors are allotted randomly to the whole plots. In a systematic design, the same procedure is followed except that the levels of crop j densities are arranged systematically in systematic 'subplots'. This design will be referred to as design II. This layout is

different from design I in that only density levels of crop j are arranged systematically unlike design I where both are arranged systematically. Model (2.12) and assumptions as for scenario I also apply here.

Scenario III:

In a split plot design, different levels of j are randomised in sub plots while in systematic design, they are systematically arranged. Spatial arrangements are allotted to main plots randomly. This design will be referred to as design III. In this setting, crop i is the main crop with fixed crop density and the interest is to investigate the effect of crop j densities on its performance i.e. to search for density combination which can maximise the responses. The probable models could be models (2.10) and (2.12).

For ease of reference, these arrangements will be referred to as R-by-C arrays where R denotes number of rows (main plots) while C denotes the number of columns (subplots to which intercrops are applied) within a row. Model (2.21) applies where Z or X and $\mathbf{e}$ are redefined to accommodate the trend and spatial dependence, respectively.

In classical analysis of randomised blocks, two underlying principles for analysis are that plot errors are independent and that global trend does not exist i.e. no spatial correlation. In systematic design, however, due to lack of randomisation, there may be spatial correlation in the data due to global trend (if the systematic arrangement falls on field trend) or local trend and extraneous trend (due to experimental materials and other environmental factors). There is a need then for modelling spatial variability. Modelling the systematic design data under different conditions of spatial variability assuming a linear mixed model situation was undertaken. The process and results are discussed in the following chapter.

Chapter 4

Modelling Systematic Design Data

4.1 Introduction

Models which encompass all the extra sources of variation that occur in the field possibly due to lack of randomisation within the main plots of the systematic designs are presented. The model that identifies the three types of variation, namely global trend (large-scale variation), local trend and also the extraneous trend which may be due to lack of randomisation is given. Generally, such models accounts for spatial correlation through modelling global trend directly and modelling local and extraneous variation through plot errors. Unlike in many previous models where 'trend' has been assumed to be due to natural variation (Martin, 1990; Cullis and Glesson, 1991; Zimmerman and Harville, 1991; Brownie *et al.*, 1993; Cressie and Hartfield, 1996), an extension to account for extraneous sources of variations has been made. Noting that either global trend, correlated errors or both may exist in the field data, the model formulations in this chapter are based on knowledge of spatial models and generalized additive models (GAMs). The subsequent sections give brief theoretical reviews and discussions of these models. The understanding of this theory is essential in understanding the modelling of intercropping experiments from systematic designs.

4.1.1 Spatial models

Historical review

Having realized that in many fields of agricultural experiments, there exists positive, or more rarely negative correlation between the errors in adjacent plots, Fisher introduced the concept of randomisation to neutralise it. This induces the assumption that errors are independent and hence avoids modelling errors. However, recently, there has been much interest in methods where experimental unit errors are modelled, especially for experiments where randomisation has not managed to reduce error correlations (Bartlett, 1978; Wilkinson *et al.*, 1983; Besag and Kempton, 1986; Williams, 1986; Martin, 1990, Zimmerman and Harville, 1991; Frensham *et al.*, 1997). This approach began with the empirical observation that variability in large agricultural field trials was inconsistent with an assumption of independence (Fairfield, 1938). There are two groups of methods that have been proposed to account for spatial dependence between experimental observations. The first group, built on the theory of *time series*, includes the *nearest neighbour* models that were initiated by the work of Papadakis (1937). The second group of methods, known as the *random field approach*, is based on the theory of regionalised variables.

The nearest neighbour class of models have been widely studied and extended to two dimensions from the original one-dimension setting (Bartlett, 1978; Cullis and Glesson, 1991) and alternative models suggested (Wilkinson *et al.*, 1983; Besag and Kempton, 1986; Glesson and Cullis, 1987). These models are generally regarded as '*trend + error*' models. The trend refers to a correlated random process represented by a low order autoregressive integrated moving average (ARIMA) process (Cullis and Glesson, 1991). In these models both trend and error are treated as random with covariance structure corresponding to that for a separable lattice process.

The random field approach is based on the theory of regionalised variables, a corner stone of geostatistics. The concept of a random field is that any response variate observed at location $\mathbf{s}$ (*centroid of experimental unit*) of experimental area $\mathcal{D}$ ($\mathcal{D} \subseteq \Re^d$, a d-dimensional Euclidean space) and generally $d = 2$ for field experiments, is a random variable $y(\mathbf{s})$. Let R_i $(i = 1, 2, \ldots, n)$ be the connected subsets i.e. plots belonging to $\mathcal{D}$ and let y_i denote the observed response on R_i and let $\mathbf{y}$ represent $n \times 1$ vector of these observations. Zimmerman and Harville (1991) have shown that $y_1, y_2, \ldots, y_n$ can be modelled in terms of a random field

$$y \equiv \{y(\mathbf{s}) : \mathbf{s} \in \mathcal{D} \subseteq \Re^2\}$$

whose members have the representation

$$y(\mathbf{s}) = m(\mathbf{s}; \beta) + \mathbf{Z}(\mathbf{s})$$

where $m(\cdot; \boldsymbol{\beta})$ is a function of a two dimensional vector (describing the cartesian coordinates of the plot centroids), $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters and $\{\mathbf{Z}(\mathbf{s}) : \mathbf{s} \in \mathcal{D}\}$ is an unobservable random field such that $E\big(\mathbf{Z}(\mathbf{s})\big) = 0$ and $\mathrm{cov}\big(\mathbf{Z}(\mathbf{s_i}), \mathbf{Z}(\mathbf{s_j})\big) = C(\boldsymbol{h}; \boldsymbol{\theta})$,where $\mathbf{h} = (\mathbf{s_i} - \mathbf{s_j})$, $\forall \mathbf{s_i}, \mathbf{s_j} \in \mathcal{D}$ and $\boldsymbol{\theta}$ is a vector of unknown parameters. The function $C(\mathbf{h})$ is called the covariogram or the *covariance function* of $\mathbf{y(s)}$. It is a *second order stationary* process i.e. depends on $\mathbf{s}_i$ and $\mathbf{s}_j$ only through the displacement vector $\mathbf{s}_i - \mathbf{s}_j$ commonly referred to as lag $\mathbf{h}$. A model for the observations y_i is obtained from that for random field y (Zimmerman and Harville, 1991).

$$y_i = y(\mathbf{s}_i) = m(\mathbf{s}_i; \beta) + \mathbf{Z}(\mathbf{s})_i$$

Another main concept in characterising random fields is a *semi-variogram* denoted $\gamma(\cdot)$ and defined by

$$var\{\mathbf{Z}(\mathbf{s}_i) - \mathbf{Z}(\mathbf{s}_j)\} = 2\gamma(\mathbf{h}; \boldsymbol{\theta})$$

$2\gamma(\cdot;\cdot)$ is called the variogram and $\gamma(\cdot;\cdot)$ is called a semi-variogram. The covariance functions and the variogram are related through the equation

$$\gamma(\mathbf{h};\boldsymbol{\theta}) = C(0) - C(\mathbf{h};\boldsymbol{\theta})$$

where $C(0) = \text{var}[y(\mathbf{s})]$.

In general, but not always, as $\mathbf{h}$ increases, the semi-variogram tends to increase from initial zero value (Journel and Huijbregts,1978). It may stop increasing beyond a certain distance and becomes more or less stable around a limit value called a '*sill*' value which is simply the *a priori* variance of the random variable. At this distance and beyond the plots are not correlated. In some cases there is a discontinuity of the variogram at the origin called a '*nugget effect*' which is due to measurement errors or white noise.

The common covariance schemes

The covariance functions that have been suggested in line with this argument are isotropic and are fitted with or without the nugget effect. Studies on uniformity trials have indicated that the most common isotropic structures of dependence between observations are the exponential, gaussian and spherical covariance functions (Samra *et al.*, 1990; Bhatti *et al.*, 1991). The equations of the principal covariance/variogram functions are presented below.

Spherical

$$\gamma(h;\boldsymbol{\theta}) = \begin{cases} \sigma^2 & h = 0 \\ \sigma^2\left(1 - \left[\frac{3h}{2a} + \frac{h^3}{2a^3}\right]\right) & 0 < h \leq a \\ 0 & h > a \end{cases}$$

Gaussian

$$\gamma(h;\boldsymbol{\theta}) = \begin{cases} \sigma^2 & h = 0 \\ \sigma^2 exp(-\frac{3h}{a})^2 & h \neq 0 \end{cases}$$

Exponential

$$\gamma(h;\boldsymbol{\theta}) = \begin{cases} 0 & h = 0 \\ \sigma^2 exp(-\frac{3h}{a}) & h \neq 0 \end{cases}$$

Nugget effect

$$\gamma(h;\boldsymbol{\theta}) = \begin{cases} \sigma^2 & h = 0 \\ 0 & h \neq 0 \end{cases}$$

where $\boldsymbol{\theta}$ is the vector of parameters (σ^2, a) and a is the practical range i.e. the distance at which the model is at 95% of the sill. By definition the autocorrelation ρ of plot observations is given by $\rho = exp(-3/a)$.

Furthermore, in spatial analysis of field experiments, some authors (Martin, 1990; Cullis and Gleeson, 1991; Zimmerman and Harvile, 1991; Verbyla and Cullis, 1992) have postulated separability of the covariance functions. The covariance separability implies that

$$C(h;\boldsymbol{\theta}) = \sigma^2 \rho_1(\mathbf{h}_1;\boldsymbol{\theta}_1)\rho_2(h_2;\boldsymbol{\theta}_2)$$

where $\rho_.(\cdot;\cdot)$ are the correlation/covariance functions, $\mathbf{h}_1 = \mathbf{s}_{i1} - \mathbf{s}_{j1}$ and $\mathbf{h}_2 = \mathbf{s}_{i2} - \mathbf{s}_{j2}$ are the lags along both axes describing the plane, i.e. along rows and columns, $\boldsymbol{\theta}$ is a vector of unknown spatial parameters.

Different from geostatistics setting are the autoregressive error (AR) structures where autocorrelation between observations decreases as the time lag between them increases. The use of an AR process for modelling error structures originated from time series modelling and have been adopted to other fields including agricultural field experiments. The adoption is based on assumption that the strength of correlation between any two plot observations is greatest for adjacent plots and diminishes as the distance between the plots increases i.e. decreases exponentially. Gilmour *et. al.* (1997) demonstrated the equivalence of AR(1) models and exponential error models, thus adopting the AR(1) models to model error structures of field experiments. The semi-variogram

of AR(1) is given as

$$\gamma(\mathbf{h}) = \sigma^2(1 - \rho^h)$$

Estimation

The estimation of the variograms can be done using a method of moments; least squares such as ordinary least squares, generalized or weighted least squares; and maximum likelihood or restricted maximum likelihood (REML), and Bayessian methods. Restricted/residual maximum likelihood (REML) estimates are said to posses less bias (Lill *et al.*, 1988; Baird and Mead, 1991). Most of the estimations in this study are based on REML.

4.1.2 Generalized Additive Models

These belongs to a general class of models called non-parametric regression models. In nonparametric regression, one is interested in estimating the mean function $E(y|\mathbf{X}) = f(\mathbf{X})$ from a set of observations $(\mathbf{X}_1, y_1), \ldots, (\mathbf{X}_n, y_n)$ without specifying a fixed functional form. $\mathbf{X}$ is a set of explanatory variables. That is the shape of the surface or curve fit to the data is determined by the data themselves, and not by the dictates of a predetermined model. Generalized additive models incorporate smoothed functions of explanatory variables into a regression-like model. Many methods are available for achieving smoothing and they include kernel-based methods, regression splines, smoothing splines, and wavelet and Fourier series expansions. Smoothing splines in particular are well known for their flexibility to fit a mean function. A model that contains both smoothed functions and parametric functions is referred to as a Semi-Parametric Additive Model (SAM). Recent research on smoothing splines have extended them to more flexible forms i.e. mixed model forms (Wang and Taylor, 1995; Wang, 1998). The GAM's can be useful in smoothing out the field trend in the data. In this work,

smoothing splines will be investigated further in relation to systematic design data. The assumption is that the trend may be represented as the sum smooth functions of plot position in the field.

In the next sections, spatial models are adopted to modelling systematic design data. Also beta-hat models are given in Section 4.4. For simplicity of reference, the spatial model adoptions given in Section 4.2 will also be referred to as general spatial models. This is to differentiate them from beta-hat models that incorporates spatial variability modelling.

4.2 Models

Three sets of models are considered in turn, namely (i) models for correlated error (ii) models for adjusting for global trend and (iii) combined models for trend and correlated errors.

4.2.1 Error models

Spatial mixed linear model

General description

Consider the scenario where correlations in observations exist but global field trend is absent. The interest then is to account for spatial variability in errors. A spatial linear mixed model is considered here for that purpose. The $\mathbf{e}$ vector of subplot errors in (2.21) consists of sub-vectors $\{e_j\}$ where e_j is the vector of plot errors for jth main plot or block. The sub-vector $\{e_j\}$ can be decomposed into components, $\boldsymbol{\varepsilon_j}$, which is a spatially dependent random error vector, and $\boldsymbol{\eta_j}$, which is an independent white noise process. Thus, for a single whole plot (j) for design II and III, the response can be modelled as

$$\mathbf{y}_j = X_j\boldsymbol{\beta}_j + Z_j\mathbf{u} + \boldsymbol{\varepsilon}_j + \boldsymbol{\eta}_j \tag{4.1}$$

where $\boldsymbol{\beta}_j$ is a t vector of treatment/fixed effects and X_j is its design matrix while $\mathbf{u}_j$ is $q \times 1$ vector of random effects with Z_j as its design matrix. It is assumed that $(\boldsymbol{u_j}, \boldsymbol{\varepsilon_j}, \boldsymbol{\eta_j})$ are pairwise independent and the plot variance (white noise process) is $\sigma^2_{\eta_j}$ and variance of spatial random error vector $\boldsymbol{\varepsilon_j}$ is $\sigma^2_j\Sigma_j(\boldsymbol{\alpha_j})$ where Σ_j is a spatial covariance matrix which is a function of parameters $\boldsymbol{\alpha_j}$ (i.e. lag and range). The error variance for the main plot j can be represented as

$$R_j = \sigma^2_j\Sigma_j(\boldsymbol{\alpha_j}) + \sigma^2_{\eta j} I_{Nj} \tag{4.2}$$

The spatial covariance matrix Σ_j can be modelled using covariance models reviewed in Section 4. 1.1.

For each jth of block of design I the model (4.1) is used to model the responses with exception that model terms now refer to block j and not whole plot j. For this design it can also be assumed that $\boldsymbol{\varepsilon_j}$ is a two dimensional process which is separable and hence, $\boldsymbol{\Sigma_j} = \boldsymbol{\Sigma_{cj}} \otimes \boldsymbol{\Sigma_{rj}}$. Where $\boldsymbol{\Sigma_{cj}}$ and $\boldsymbol{\Sigma_{rj}}$ are correlation matrices for columns and rows respectively.

Thus for the complete experiment the following gives the complete model (2.21)

$$\mathbf{y} = X\boldsymbol{\beta} + Z\mathbf{u} + \boldsymbol{\varepsilon} + \boldsymbol{\eta} \tag{4.3}$$

where $\boldsymbol{\beta}$ is a vector of fixed effects whose subsets are $\boldsymbol{\beta_j}$ and X is its design matrix while $\mathbf{u}$ has subsets $\mathbf{u}_j$ with Z as its design matrix consisting of Z_j's. The sub-vectors $\boldsymbol{\eta_j}$ and sub-vectors $\boldsymbol{\varepsilon_j}$ constitute $\boldsymbol{\eta}$ and $\boldsymbol{\varepsilon}$, respectively. All Z_j's are independent since observations from different main plots or blocks are independent. It is assumed that $(\boldsymbol{u}, \boldsymbol{\varepsilon}, \boldsymbol{\eta})$ are pairwise independent and their joint distribution is Gaussian with mean

zero and variance-covariance matrix

$$\begin{bmatrix} \boldsymbol{u} \\ \boldsymbol{\varepsilon} \\ \boldsymbol{\eta} \end{bmatrix} \sim \mathrm{N}\left\{\mathbf{0},\ \sigma^2 \begin{pmatrix} G_{(\gamma)} & 0 & 0 \\ 0 & \sum(\boldsymbol{\alpha}) & 0 \\ 0 & 0 & \boldsymbol{\psi} I \end{pmatrix}\right\}$$

where γ is a vector of variance components related to possible sub-vectors in $\mathbf{u}$ and $\boldsymbol{\Sigma}$ is a spatial covariance matrix which is a function of parameters $\boldsymbol{\alpha}$ (h and a). The matrix G is a function of a vector of unknown parameters γ and may or may not be completely general i.e. completely unstructured. The notation $G_{(\gamma)}$ stresses the dependence of G on γ. The matrix $\boldsymbol{\psi} I$ is associated with a plot variance, σ_η^2, with $\psi = \sigma_\eta^2/\sigma^2$. The marginal distribution of $\mathbf{y}$ is assumed to be normal and is given as

$$\mathbf{y} \sim N\Big(X\boldsymbol{\beta},\ \sigma^2\big(ZGZ' + R(\phi)\big)\Big)$$

where $R(\phi) = \Sigma(\boldsymbol{\alpha}) + \boldsymbol{\psi} I$ and $\phi = (\boldsymbol{\alpha}', \boldsymbol{\psi})'$. Let $H = ZGZ' + R(\phi)$ and $R = R(\phi)$, then R is error variance which models the structure of common covariance and residual variance (Littell *et al.*, 1996) and is a block diagonal matrix. Matrix H is a block diagonal matrix since the observations from different main plots or blocks are assumed uncorrelated. In all the discussions that are to follow an assumption that H^{-1} exists is made. The covariance models used in geostatistics such as spherical, gaussian and exponential, and their directional forms are suggested for modelling $\boldsymbol{\varepsilon}$ (spatial dependence). Also AR(1), double first order autoregressive model (AR(1)×AR(1)) and first order autoregressive moving average (ARMA (1, 1)) are suggested. There are more other covariance models but the interest in the present study is to discuss those that can easily be implemented in the available commercial statistical software.

The modelling of spatial variability using the above model will result in increased precision in modelling of data from systematic designs. The advantage of increased precision through modelling spatial variability has been demonstrated by many authors (Gilmour *et. al.*, 1997).

Estimation

The estimation of fixed effects and random effects in (4.3) are obtained by solving the mixed model equations (MMEs) derived by Henderson (1984). For the spatial linear mixed model (4.3), the joint probability density function of $\mathbf{y}$ and $\mathbf{u}$ is given by

$$\begin{aligned} f(\mathbf{y},\mathbf{u}|\boldsymbol{\beta},\boldsymbol{\theta},\sigma^2) &= f(\mathbf{y}|\mathbf{u})f(\mathbf{u}) \\ &= \frac{\exp\left\{-\frac{1}{2\sigma^2}\left[(\mathbf{y}-X\boldsymbol{\beta}-Z\mathbf{u})'R^{-1}(\mathbf{y}-X\boldsymbol{\beta}-Z\mathbf{u})+\mathbf{u}'G^{-1}\mathbf{u}\right]\right\}}{(2\pi)^{0.5v}\{det(R)\}^{0.5}\{det(G)\}^{0.5}} \end{aligned}$$

where $v = n - t$. Setting the partial derivatives of this function with respect to $\boldsymbol{\beta}$ and $\mathbf{u}$ to zero, the following equations are obtained

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z+G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} X'R^{-1}\mathbf{y} \\ Z'R^{-1}\mathbf{y} \end{bmatrix}. \tag{4.4}$$

These equations are called mixed model equations. Note that the MMEs are extended normal equations. Without G^{-1} in the lower right hand sub-matrix, i.e. if G^{-1} is zero, the random effects are estimated as though they were fixed effects. Also observe that in the above equations, it is assumed that G is non-singular. In the case where it is singular, then modification by setting the elements of $\mathbf{u}$ corresponding to singular portion of G equal to zero is effected (Henderson, 1984). The solution to the equations (4.4) yields the generalized unbiased estimates (GLEs) of fixed effects and Best linear unbiased predictors (BLUP) of random effects as,

$$\hat{\boldsymbol{\beta}} = (X'H^{-1}X)^{-1}X'H^{-1}\mathbf{y} \tag{4.5}$$

$$\hat{\mathbf{u}} = GZ'P\mathbf{y} \tag{4.6}$$

$$\text{where } P = H^{-1} - H^{-1}X(X'H^{-1}X)^{-1}X'H^{-1}$$

The variance of the estimates of fixed effects is given by

$$Var(\hat{\boldsymbol{\beta}}) = (X'H^{-1}X)^{-1} \tag{4.7}$$

while the joint variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ is given by

$$C = \begin{pmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{pmatrix}^{-} \sigma^2 \tag{4.8}$$

where $(^-)$ denotes a generalised inverse. Note that both variances (4.7) and (4.8) give the same conclusions about the fixed effects (Robinson, 1991). The estimates of G ($\hat{G}$) and R ($\hat{R}$) are used whenever they are unknown. In the present study the estimates of G and R are estimated using restricted/residual likelihood method (REML). Although, there are many other methods for estimation, REML estimates are said to possess low bias (Lill *et al.*, 1988; Baird and Mead, 1991).

The vector of variance components $\boldsymbol{\theta}$ (γ, ϕ) are obtained as REML estimates. Note that H can also be written as $H(\boldsymbol{\theta})$ to stress the dependence of the covariance matrix on $\boldsymbol{\theta}$. Assuming multivariate normal distribution, the REML likelihood can be given by the expression below (Patterson and Thompson, 1971).

$$\ell_R = -\frac{1}{2}(\ln|H| + ln(|X'X|) + \ln|X'H^{-1}X| + \mathbf{y}'P\mathbf{y}$$

Minimising the above log-likelihood requires iterative methods such as Newton-Raphson or Fisher scoring method. For numerical and computational efficiency both the first and second derivatives of ℓ_R are obtained. According to McCulloch and Searle (2000) the expectations of the second derivatives form the information matrix of the covariance parameters, i.e.

$$\mathcal{J}_{(\theta\,\sigma^2)} = -E\begin{bmatrix} \frac{\partial^2\ell_R}{\partial\theta_i\partial\theta_{i'}} & \frac{\partial^2\ell_R}{\partial\theta_i\partial\sigma^2} \\ \left(\frac{\partial^2\ell_R}{\partial\theta_i\partial\sigma^2}\right)' & \frac{\partial^2\ell_R}{\partial\sigma^4} \end{bmatrix} \tag{4.9}$$

The inverse of this information matrix gives the asymptotic covariance matrix of the covariance parameter estimates. This provides a basis of testing the fit of different

covariance models to the data.

The REML algorithm (e.g. as implemented in Genstat or PROC MIXED in SAS (SAS Institute, 2001)) first obtains initial estimates of the variance components from an ordinary least squares fit ignoring the random effects. These estimates are then used to calculate the estimates of $\boldsymbol{\beta}$ and $\mathbf{u}$ by inverting the mixed model equations. Using $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$, the first derivatives of ℓ_R and the elements of matrix $\mathcal{J}_{(\theta\sigma^2)}$ are formed. Then using Fisher scoring (in Genstat) or Ridge stabilised Newton-Raphson (as in PROC MIXED) method, the estimates of the variance components are updated. This process is iterated till convergence of variance parameter estimates. Relative Hessian convergence criterion with a default tolerance number of 1×10^{-8} is used to check for convergence (SAS Institute Inc., 2001). This criterion makes use of first derivative (say g_k) of the objective or likelihood function at kth iteration and the inverse of the second derivative/Hessian matrix (H_k^{-1}). This criterion is defined by $\frac{g_k' H_k^{-1} g_k}{|f_k|} \leq 1 \times 10^{-8}$ where $|f_k|$ is the value of the objective function.

4.2.2 Trend models

General description

Although, the blocking principle of experimental design is applied in the systematic designs, the chances that treatments in a systematic design will fall on a global trend are higher compared to randomised blocks. The global trend may exist within the main plot and thus there is a need to account for it to remove biases in treatment effects and error variance estimates. This is also referred to as *trend analysis*. Although the presence of global trend cause a spatial correlation in the observations, the errors can be assumed to be uncorrelated. The spatial trend is modelled as a function of plot positions. Consider a rectangular $q \times p$ layout of plots, with row position indexed by R_i, where $i = 1, \ldots, q$ and column indexed by C_j, where $j = 1, \ldots, p$ or more compactly

denote plot position by vector $\mathbf{t}$ for n total number of plots in the experiment. Two models, namely a spatial mixed model with polynomial mean function in the field trend and a GAM are presented.

Linear mixed model

General

A linear mixed model is extended to incorporate a trend (T_{ij}), which is assumed to be a polynomial function of R_i and C_j or can alternatively be represented using orthogonal polynomials in R_i and C_j. For instance consider the trend modelled as a quadratic function as

$$T_{ij} = b_1 R_i + b_2 C_j + b_3 R_i^2 + b_4 C_j^2 + b_5 R_i C_j \tag{4.10}$$

where b_i's are regression parameters. As explained by Kirk *et al.* (1980), fitting a polynomial response surface corresponds to partitioning out of error the systematic component of heterogeneity and the estimates of precision are based on the remaining random component. The trend, T_{ij}, can be assumed to have fixed and random components. T_{ij} terms are included in the design matrix X part of model (2.21) to give;

$$\mathbf{y} = X_* \boldsymbol{\beta}_* + Z\mathbf{u} + \mathbf{e} \tag{4.11}$$

where $\boldsymbol{\beta}_*$ is a vector corresponding to fixed effects including treatment and trend effects and X_* is design matrix including both treatment effects and trend effects, and e is a vector of residuals that are independent $e \sim (0, \sigma_\epsilon^2 I)$.

For this model (4.11), the variability among plots within the same block/main plot is made up of two components, one due to the trend model and one due to plot-to-plot variation. The entry of the trend in the mixed model induces some extra variability into the traditional random component, σ^2. Thus σ^2 can be partitioned into two components i.e.

$$\hat{\sigma}^2 = \text{MSTrend} + \hat{\sigma}_\epsilon^2 \approx \hat{\sigma}_T^2 + \hat{\sigma}_\epsilon^2$$

where

$$\text{MSTrend} = \frac{1}{rt}\sum_{k}^{r}\sum_{l}^{t}\left(T_{hl} - \bar{T}_{h.}\right)^2$$

where T_{hl} is the trend value in lth $(i, j$th) plot in the hth block or main plot. Thus, the variance of $\mathbf{y}$ is given by

$$\begin{aligned} \text{var}(\mathbf{y}) &= ZGZ' + (\sigma_T^2 + \sigma_\epsilon^2)I \\ &= ZGZ' + R \;=\; V \end{aligned}$$

Estimation

Model (4.11) is a linear mixed model and thus the estimates of $\boldsymbol{\beta}$ and $\mathbf{u}$ are obtained as solutions to the usual mixed model equations. The estimates are

$$\begin{aligned} \hat{\boldsymbol{\beta}}_* &= (X_*'V^{-1}X_*)^{-1}X_*'V^{-1}\mathbf{y} \\ \hat{\mathbf{u}} &= GZ'P\mathbf{y} \\ &= GZ'V^{-1}(\mathbf{y} - X_*\hat{\boldsymbol{\beta}}) \end{aligned}$$

$$\text{where } P = V^{-1} - V^{-1}X_*\left(X_*'V^{-1}X_*\right)^{-1}X_*'V^{-1}$$

The estimates of variance components (σ^2, γ) are obtained as REML estimates.

Based on the above trend model, simulations were done to investigate the effect of excluding modelling the effect of replicates/blocks. It was observed from 600 realizations that the error variance and standard errors for treatment contrasts from a model with blocks excluded are twice as those from the model with blocks included (Table V; Appendix C). The realizations were generated by assuming twelve treatment effects and four blocks of RCBD. In each simulation a normally distributed variates with varying means (12 different mean values) and variance of 1 was generated using GenStat rendom number generator. The 12 different means were randomly chosen in the range of 100 and 130 using GenStat random number generator to represent the 12 treatment effects. Assuming the 12 treatments were arranged in a single row (i.e. a block consists of a single

row) a quadratic trend effect ($T_{ij} = 1.89 + 0.4R_{(i)} - 0.18R^2_{(i)}$) in rows was superimposed on the generated variates where $R_{(i)}$ is the plot position number - average position number of plots. The code for simulation is included in Appendix B. The choice was therefore to include these effects in the model.

Semi-parametric mixed model

General description

In this approach, the trend is represented using the 'data-driven' smoothing method of Generalized additive models. The generalized additive models (GAM's) were introduced by Hastie and Tibshirani (1990) and assume that the trend may be represented as the sum of smooth functions, without specifying a fixed functional form. This is in contrast with modelling trend using a polynomial mean function where a polynomial has to be specified. Suppose data y_i is observed on plot position t_i. Then according to Wahba (2000), smoothing is a penalized least squares problem in which the function f is chosen as a solution to the penalized objective function

$$\arg \min_f \frac{1}{n}\sum_{i=1}^{n}[y_i - f(t_i)]^2 + \lambda \int [f^m(t)]^2 dt \tag{4.12}$$

where $n = rt$ for t treatments and r replicates. Equivalently in matrix form as

$$\arg \min_{f \in \mathcal{H}} \left\{ \frac{1}{n}(\mathbf{y} - \mathbf{f})'(\mathbf{y} - \mathbf{f}) + \lambda ||P_1 f||^2 \right\} \tag{4.13}$$

where $f \in \mathcal{H} = \mathcal{H}_0 \bigoplus \mathcal{H}_1$ with $\mathcal{H}_0$ a finite dimensional space containing basis functions that are not to be penalized and $\mathcal{H}_1$ is a Sobolev Hilbert space of functions with $m-1$ continuous derivatives ($\int_0^1 [f^m(t)]^2$). Thus, P_1 is the orthogonal projection of f onto $\mathcal{H}_1$ in $\mathcal{H}$ i.e. a penalty to the departure of f from the space $\mathcal{H}_0$, i.e. a penalty to the roughness of f. The parameter λ determines the smoothness and goodness of fit of the function f. If $m = 2$ then a solution to this minimisation problem is a natural cubic smoothing spline.

In the present study the interest is in smoothing the possible field trend introduced in the observations due to systematic arrangement of treatments. This is done by using plot positions as variates. That is the model to be used is of the form

$$\mathbf{y} = X\boldsymbol{\beta} + f_1 + \ldots + f_q + \mathbf{e} \tag{4.14}$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients and X is the design matrix, f_i $(i = 1, 2, \ldots, q)$ is the smooth term which is a function of covariate t_i and $\mathbf{e}$ is a vector of independent errors with common variance σ^2. This is called the semi-parametric additive model (SAM). The spline function f represents the sum of the smooth terms f_i's. Following Green (1987), the spline function f is of the form

$$\mathbf{f} = X_s\boldsymbol{\beta}_s + N_s\mathbf{u}_s$$

where $\mathbf{N_s} = \boldsymbol{\nabla}(\boldsymbol{\nabla}'\boldsymbol{\nabla})^{-1}$ and $\boldsymbol{\nabla}$ is an $n \times r(t-2)$ matrix satisfying $\boldsymbol{\nabla}'X_s = 0$ and X_s is a block diagonal matrix with each block having the entries $[1, x_s]$ where x_s is a $t \times 1$ vector of distinct t_i's and $\mathbf{1}$ is a $t \times 1$ vector of ones. Using Verbyla *et. al.* (1999) definitions, $\boldsymbol{\nabla}$ is a block diagonal matrix with each ith block $(i = 1, \ldots, r)$ being a $t \times (t-2)$ banded matrix. This matrix is also known as the difference matrix that defines smoothing procedure the elements of which are

$$\boldsymbol{\nabla}_i = \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ -2 & 1 & 0 & 0 & \ldots & 0 \\ 1 & -2 & 1 & 0 & \ldots & 0 \\ 0 & 1 & -2 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 \end{pmatrix} \tag{4.15}$$

Thus model (4.14) becomes

$$y = X\boldsymbol{\beta} + X_s\boldsymbol{\beta}_s + N_s\mathbf{u}_s + \mathbf{e} \tag{4.16}$$

where $\boldsymbol{\beta}$ is a vector of fixed effects (treatments and block effects) and X its design matrix while the trend is represented by $X_s\boldsymbol{\beta}_s+N_s\mathbf{u}_s$. In the present case, $X_s\boldsymbol{\beta}_s$ represents a fixed linear trend along the main plot or block and $N_s\mathbf{u}_s$ represents random non-linear trends, with $\mathbf{u}_s$ as a set of correlated random normal deviates with mean 0 and variance-covariance of $\sigma_s^2\Lambda_s$ or $\sigma^2/\lambda\Lambda_s$. The residual $\mathbf{e}$ is an *iid* variable with variance σ^2. The diagonal blocks of matrix Λ_s associated with each $\boldsymbol{\nabla_i}$ defined above are $t \times (t-2)$ matrices given by

$$\Lambda_{si} = \begin{pmatrix} 4 & 1 & 0 & 0 & \dots & 0 \\ 1 & 4 & 1 & 0 & \dots & 0 \\ 0 & 1 & 4 & 1 & \dots & 0 \\ 0 & 0 & 1 & 4 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 4 \end{pmatrix} \tag{4.17}$$

In this set up equation (4.13) then becomes

$$\arg \min_{f \in \mathcal{H}} \left\{ \frac{1}{n}(\mathbf{y} - X\boldsymbol{\beta} \text{ - } \mathbf{f})'(\mathbf{y} - X\boldsymbol{\beta} - \mathbf{f}) + \lambda ||P_1 f||^2 \right\} \tag{4.18}$$

Estimation

Denote variance components $\boldsymbol{\theta}$ (σ^2, σ_s^2) and $\tau = 1/\lambda$ as parameters to be estimated. Then by treating the above semi-parametric model (4.16) as a mixed-effects model, $\boldsymbol{\theta}$ and τ are obtained as REML estimates. In this case, τ is treated as an extra variance component in a linear mixed model. The parameter τ is equal to the ratio σ_s^2/σ^2 (Wang, 1998; Verbyla *et al.*, 1999). Thus the variance components (σ^2, σ_g^2, λ) are obtained by minimising the residual log likelihood.

$$\ell_R(\tau, \theta, y) = -\frac{1}{2}\{ \ln|V| - \ln|X_* V^{-1} X_*| \} - \mathbf{y}' P \mathbf{y} \tag{4.19}$$

where $X_* = [X : X_s]$, V has block diagonal entries $V_i = \sigma^2(I + \lambda^{-1}N_{si}\Lambda_{si}N'_{si})$,

$$\text{and } P = V^{-1} - V^{-1}X_*(X'_*V^{-1}X_*)^{-1}X'_*V^{-1}$$

The minimisation is an iterative process that makes use of first and second derivatives of ℓ_R.

The values of λ vary from 0 (interpolation) to ∞ (linear trend). Separate values of λ can be fitted for different blocks/main plots as the strength/nature of global trend may be different in different blocks/main plots. A log likelihood ratio test ($\ell = -2(\ell_0 - \ell_1)$) can be used to test for equality of different λ's. Where ℓ_0 and ℓ_1 are values of the residual log likelihood under H_0 of equality of λ's and H_1 of at least one of the λ's different. However, in this study common λ would be assumed known.

The estimate of treatment effects adjusted for the trend obtained by solving the mixed model equations is

$$\hat{\boldsymbol{\beta}} = (X'(I - M)X)^{-1}X'(I - M)\mathbf{y} = A\mathbf{y}. \tag{4.20}$$

M is called a *centered hat-matrix* or smoothing matrix as defined by Durbán *et al.* (1997) and it is calculated as M $n \times n$ block diagonal matrix with block diagonal entries $M_i = M_i(\lambda) = \left(I + \lambda\boldsymbol{\nabla}_i\Lambda_{si}^{-1}\boldsymbol{\nabla}_i'\right)^{-1} - \mathbf{1}\mathbf{1}'/t$. It is a hat-matrix of the smoothing terms (additive part) of the model. The estimate of the trend is $M(y - X\hat{\boldsymbol{\beta}})$ and the variance of treatment estimates is $Var(\hat{\boldsymbol{\beta}}) = AA'\sigma^2$. The estimate of σ^2 above is identical to conventional estimator $\hat{\sigma}^2 = \sum_{i=1}^{n}\left(y_i - f(t_i)\right)^2/tr(M)$ for an additive model (Green and Silverman, 1994). Thus, tr(M) also measures smoothness of f. It should be noted that λ and tr(M) are inversely related.

4.2.3 Joint models for trend and error

It is very possible that both strong field trend and correlation in errors may exist in data observed from a systematic design. It has been shown that modelling both present is more efficient than modelling only one of them (Brown and Gumpertz, 1995).

Although, it has been shown that fitting a covariance function 'soaks up' most spatial heterogeneity (Zimmerman and Harville, 1991), modelling both is still most effective (Brownie *et al.*, 1993). Modelling error terms using covariance function and trend terms using a cubic smoothing spline or a polynomial mean function would be effective. A spatial linear mixed model with trend accounted for using a polynomial mean function and a spatial semi-parametric mixed model where the trend is modelled using a cubic smoothing spline are suggested for modelling systematic design data. A brief introduction to these models is give below. The estimation in these models is generalised from those models in Section 3.5.2.

Spatial linear mixed model with trend incorporated

This approach models large-scale spatial trends through fixed effect polynomial terms and allows for small scale and extraneous variation through correlation between neighbouring plots. This is correlated error modification of the trend analysis. That is

$$\mathbf{y} = X_*\boldsymbol{\beta}_* + Z\mathbf{u} + \boldsymbol{\varepsilon} + \boldsymbol{\eta} \tag{4.21}$$

This model is basically like model (4.11) except that $\mathbf{e}$ is decomposed into $\boldsymbol{\varepsilon}+\boldsymbol{\eta}$ to allow for the modelling of spatial variability. However, if block effects are assumed random then model (4.21) can be rewritten as

$$y = X\boldsymbol{\beta} + Z\mathbf{u} + \boldsymbol{\varepsilon} + \boldsymbol{\eta} \tag{4.22}$$

All the estimation and prediction procedures for the above models are as for those given for model (4.11), with R replaced by

$$R = \sigma^2\big(\Sigma(\boldsymbol{\alpha}) + \psi I\big)$$

Spatial semi-parametric mixed model

Extensions to model (4.16) to allow for correlation in $\mathbf{e}$ can be defined by decomposing $\mathbf{e} \sim N(0,\, \sigma^2 R)$ and the extended model is given as

$$y = X\boldsymbol{\beta} + X_s\boldsymbol{\beta}_s + N_s\mathbf{u}_s + \boldsymbol{\varepsilon} + \boldsymbol{\eta} \tag{4.23}$$

The local trend (spatial correlation in errors) is modelled through $\boldsymbol{\eta}$ and white noise is modelled through $\boldsymbol{\varepsilon}$ with common variance σ^2. In this setting, the marginal distribution of $\mathbf{y}$ is

$$\mathbf{y} \sim N\big(X\boldsymbol{\beta} + X_s\boldsymbol{\beta_s},\, \sigma^2 H(\boldsymbol{\theta})\big)$$

where $H(\boldsymbol{\theta}) = R + \sigma^2(\lambda^{-1}N_s\boldsymbol{\Lambda}_s N_s')$ and $R = R(\phi) = \Sigma(\boldsymbol{\alpha}) + \psi I,\ \phi = (\boldsymbol{\alpha}',\, \psi)'$

The estimation of the vector of smoothing parameters $\boldsymbol{\lambda}$ and other components of variance is achieved through minimising the residual log likelihood in (4.19) with V replaced by $H(\boldsymbol{\theta})$ and $M_i = \big(I^{-1} + \lambda\boldsymbol{\nabla}_i\Lambda_{si}^{-1}\boldsymbol{\nabla}'_i\big)^{-1}R^{-1}$.

The preceding discussions illustrate how data from systematic designs can be handled using linear mixed model form of general linear models. The ease of modelling spatial variability i.e. global trend and spatial error structure is shown by use of spatial linear mixed models. In brief, analysis of data with spatial variability such as systematic design data is easily implementable.

4.2.4 Inference on linear models

Model selection

All the models considered in this chapter are linear models. In fact all of them are linear mixed models. This means that the usual available tools can be used to test for adequacy of these models and testing hypotheses or setting confidence intervals (CIs).

The basic idea of model selection is that all models fitted have fixed effects which may be fitted with or without the random effect terms or with different covariance models. The interest is in selecting a model which provides the 'best' estimation of fixed effects. The following tools are used to check for adequacy.

1. Residual log likelihood ratio test (LRT) or deviance. Consider the parameter vector $\psi \equiv \{\theta, \beta\}$ that takes on its values in the parameter saturated/maximal set Ψ (where all possible parameters as there are data points are included in the model). Let Ψ_0 be the subset of Ψ corresponding to the parameter set of a restricted model. Then the LRT for $H_0 : \psi \in \Psi_0$ is defined as $-2ln\Lambda = -2(\ell(\psi \in \Psi)$ - $\ell(\psi \in \Psi_0)$.

2. Akaike Information Criterion (AIC) and Schwarz's Bayesian Criterion (BIC). The larger the values of the these two, the better the model. They are computed for m covariance parameters as

$$AIC = \ell(\psi; \mathbf{y}) - m \text{ and } BIC = \ell(\psi; \mathbf{y}) - \frac{m}{2} ln(n-p)$$

where m is the number of covariance parameters and p is the number of fixed parameters. In the present setup, the maximal value of p is 7 while that of m is 4. The term $\ell(\psi; \mathbf{y})$ refers to the residual loglikelihood of a model fitted with only fixed terms. The AIC and BIC measures the penalization due to incorporating a covariance model with m parameters.

3. Residual Mean Square Error of covariance parameter estimates and use of residual plots (such as VPLOT in Genstat) and variograms plots provide an idea about the adequacy of the model.

4. Asymptotic covariance matrix of estimates. The asymptotic covariance matrix of covariance parameter estimates is also helpful in choosing the most appropriate covariance model. This matrix is approximated from the Fisher Information matrix. A rule of good sense is to choose the adjusted covariance model (sample variogram) with the smallest covariances or standard errors. This criterion is used along with other criteria.

Note that, during model selection, the inclusion of extra fixed term can be monitored by changes in deviance or LRT.

Model inference

The estimates of fixed effects $\hat{\boldsymbol{\beta}}$ are asymptotically normally distributed. Since each $\hat{\boldsymbol{\beta}}$ has its associated variance, the approximate confidence intervals can be constructed for each individual parameter. The t-test can be used to test for significance. However, some of these models have more than one source of variation and this implies that the ratio of effect estimates to their standard errors does not follow a t-distribution (Levin, 1999) since these variances are sums of various stratum variances. Thus, Satterthwaite's formula is used to calculate approximate degrees of freedom for t-tests.

To make inferences about treatments, observe that the multivariate normal distribution assumption on $\mathbf{y}$ implies $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) \sim N(\boldsymbol{\beta}, (X'H(\boldsymbol{\theta})^{-1}X)^{-1})$ on which inference can be based. In practice the covariance parameters $\boldsymbol{\theta}$ are not known and is replaced by its estimate. Consider an estimable linear combinations or contrasts of the form $L\boldsymbol{\beta}$, and test the hypothesis that H_0: $L\boldsymbol{\beta} = 0$ against the general alternative. Then two available

test are the Wald statistic (T) and approximate F-test (Littel *et al.*, 1996)given by

$$T = \hat{\boldsymbol{\beta}}' L' \left(L (X' H(\hat{\boldsymbol{\theta}})^{-1} X)^{-1} L' \right)^{-1} L \hat{\boldsymbol{\beta}}$$

$$F = \frac{\hat{\boldsymbol{\beta}}' L' \left(L (X' H(\hat{\boldsymbol{\theta}})^{-1} X)^{-1} L' \right)^{-1} L \hat{\boldsymbol{\beta}}}{\text{rank}(L)} \tag{4.24}$$

F is approximately distributed as Fisher's F distribution with $rank(L)$ as numerator degrees of freedom and residual degrees of freedom (v) as denominator i.e $F_{[\text{rank}(L),v,\lambda_L]}$ where λ_L is a non-centrality parameter. Under H_0, T is distributed as $\chi^2_{\text{rank(L)}}$. In SAS PROC MIXED, T is given as a Type III statistic. The use of ESTIMATE and CONTRAST statements provide a way of testing different hypotheses. Rank(L) is usually approximated using Satterthwaite or Containment method. However, for small samples the asymptotic distribution of the estimates may be inadequate. Kenward and Roger (1997) suggested the use of a modified Wald statistic together with an F approximation for the sampling distribution. This statistic uses an adjusted estimator of covariance matrix of treatments and an adjusted residual degrees of freedom. This procedure is referred to as Kenward-Roger method.

4.3 Validation of suggested models

Various methods exist for simulating data to be used in model validation. In this study the Monte Carlo simulations technique was used. The method employed Cholesky decomposition of the covariance structure (Cressie, 1991). To represent what possibly would happen in the field, seven covariance functions were used. This is because it is not possible to know the exact form of spatial variability until the data analysis stage, thus the need to evaluate many covariance models. These included exponential, Gaussian, spherical, AR(1), AR(1)× AR(1) and ARMA (1,1). The values of a practical range parameter a (definition of a is given in Section 4.1.1) used included 2, 4, 6 and 8 (chosen to cover a wide range of correlation). Three factors i.e. plant spatial arrangement (with

1 to 3 levels), plant density of crop A (with 1 to 4 levels) and plant density of crop B (with 6 to 12 levels) were assumed in the simulations.

The simulations were based on the designs I, II and III discussed in Section 3.3 and assumed no treatment effect. The level of spatial heterogeneity used was chosen to be a representative of results from Brownie and Gumpertz (1997) and that generated from two field data sets from intercropping experiments from Uganda. Data were simulated for each of the mentioned covariance functions with and with out the global trend (regular or irregular) included and in some simulations only the presence of the global trend is assumed. For each setting 600 Monte Carlo samples were simulated. Simulations were done in SAS (other statistical softwares used were Gauss and SPLUS) and model fitting was done using PROC MIXED and PROC GAM in SAS, (some implementations were done using lme and slr in SPLUS, and REML and GAM in Genstat procedures). The computer simulations are used mainly because they offer a valuable and feasible alternative to empirical modelling and in particular they enable the exploitation of many spatial models and trends.

Simulation procedure

Suppose the levels of spatial arrangements are denoted as p, levels of crop A densities are denoted as a and those of crop B as b. This gives a total of abp treatments. Consider simulating observations for yields of crop A or B.

1. Let the vector $(\mathbf{f}^n)$ of dimension $abp \times 1$ represent the variates for nth simulation and let $m_{ijk}(k = 1, \ldots, p;\ i = 1, \ldots, a;\ j = 1, \ldots, b)$ represent random mean value for (ijk)th treatment randomly chosen from a given range of numbers for each nth simulation. Values of f_{ijk} are generated as normally distributed deviates with means m_{ijk} and variance of 1 for each nth simulation. The vector $\mathbf{f}^n$ of f_{ijk} represents the effect of treatments on crop yields.

2. Seven covariance matrices (H), namely exponential, Gaussian, spherical, AR(1), AR(1)× AR(1) and ARMA (1,1) were generated in SAS using a program written with SAS MACRO and PROC IML. The different global trends were also generated using a program written in SAS IML.

3. The vector $\mathbf{x}$ of $abp \times 1$ distributed variates with mean zero and variance 1 was generated for the purpose of establishing error structure. The error vector $\mathbf{e}$ was formed by $\mathbf{e} = L\mathbf{x}$ (where L is the lower triangular Cholesky factor of H). The general equation used in simulations is $\mathbf{y} = \mathbf{f^n} + \mathbf{e}^*$ where $\mathbf{y}$ is the formed yield data vector and $\mathbf{e}^*$ is taken to be correlated or uncorrelated error vector depending of whether $\mathbf{e}$ or $\mathbf{x}$ is chosen for simulations.

4. The design matrix for the designs I, II and III were generated from GenStat statistical software, each design containing *abp* plots. For each design, the design matrix was combined with the yield data formed in step 3 i.e. yields were allocated to the plots. This constituted an input file for the case of design II and III. For design I step 5 was implemented.

5. Notice that the data generated in step 4 above when viewed under design I give only spatial variability along rows. To incorporate spatial variability along columns the column factor levels in each replicate are sorted in ascending order and an appropriate $\mathbf{e}^*$ vector added.

6. For the realizations that involved inclusion of the global trend, the values of trend were generated and added to the data to form a complete input data file. A vector $\mathbf{t}$ of quadratic global trend simulated using different functions depending on number of plots per row was used in the simulations. For example using results from the Uganda data, for six plots in a row $\mathbf{t}$ was generated by $\mathbf{t} = 1.85 + 0.3R_{(i)} - 0.15R^2_{(i)}$. Where $R_{(i)}$ = subplot position number - average position

number of subplots .

The simulation procedures and GenStat and SAS codes are included in appendix E.

Simulation Results

From the construction point of view, design III contains some properties similar to design II in reference to spatial variability. In other words the results pertaining to design II concerning modelling spatial variability also applies to design III. Therefore, only results for design II are presented here since these results were the same as those from design III. The analysis of simulated data was done using PROC GAM and PROC MIXED in SAS, and REML in GenStat. Except for the spherical model where ARMA(1, 1) model was also fitted to the data simulated using spherical model, models were fitted to the data which was simulated using them. Simulation results of the spherical covariance model are given in Table 4.1 and Table 4.2. In Table 4.1 the trend was fitted using a polynomial in plots while in Table 4.2 the trend was fitted using the cubic smoothing splines. Results for Gaussian and exponential covariances are given in Tables 4.3 and 4.4 respectively. Results for AR and ARMA models are presented in Table 4.5.

Discussion of simulation Results

The mean are the means of the given parameters from 600 Monte Carlo simulations/samples. The following observations are based on results given in Tables 4.1 to 4.4.

1. Except for processes characterised by a range parameter larger than half the length of the main plot, the covariance parameters are well estimated. The biases in the parameters are less than 10%. The standard deviations (SD) for the parameter estimates are generally low.

2. Smoothing splines seem to have performed better than polynomial mean functions

Table 4.1: REML Covariance parameters for spherical errors: trend model using polynomial mean function

Array	Theoretical values		Estimated values					
	Parameter		*Mean (no trend)*			*Mean (+ trend)*		
	σ^2	a	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)
4 by 6	1	2	0.96	2.01	0.154	0.99	2.31	0.194
4 by 6	1	4	0.98	4.18	0.239	1.01	3.93	0.214
4 by 6	1	6	0.92	5.96	0.425	0.98	6.24	0.315
4 by 8	1	2	1.01	2.08	0.530	1.01	1.96	0.630
4 by 8	1	4	0.88	3.95	0.380	1.00	3.96	0.231
4 by 8	1	6	0.96	6.01	0.231	0.92	5.98	0.411
8 by 12	1	4	0.92	4.30	0.330	0.99	3.89	0.222
8 by 12	1	6	0.98	5.88	0.322	1.01	5.97	0.351
8 by 12	1	8	0.98	8.09	0.04	0.89	7.69	0.248
8 by 8	1	4	0.95	4.20	0.950	0.97	3.68	0.835
8 by 8	1	6	1.02	5.82	0.450	1.10	6.01	0.933
8 by 8	1	8	0.98	8.12	0.431	0.98	7.97	0.881

in accounting for the trend due to their flexibility. Although this difference is not apparent, in practice where there is no knowledge about the nature of the trend, smoothing splines are likely to fit the trend more efficiently than the polynomial mean function. In some cases where data were simulated assuming a quadratic trend, smoothing splines fitted a linear spline in the plots, the reason being that accounting for spatial variation through correlation structure tend to 'soak up' spatial heterogeneity (Brownie and Gumpertz, 1997) due to a trend reducing to

Table 4.2: REML Covariance parameters for spherical errors: Using smoothing splines

Array	Theoretical values		Estimated values					
	Parameter		*Mean (no trend)*			*Mean (+ trend)*		
	σ^2	a	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)
4 by 6	1	2	0.96	2.01	0.502	0.99	2.31	0.379
4 by 6	1	4	0.98	4.16	0.314	1.00	3.86	0.169
4 by 6	1	6	1.22	5.36	0.502	1.08	5.69	0.241
4 by 8	1	2	1.01	2.08	0.861	1.01	1.98	0.308
4 by 8	1	4	0.88	3.95	0.563	1.01	3.86	0.397
4 by 8	1	6	0.96	6.01	1.010	0.97	5.98	0.103
8 by 12	1	4	0.92	4.30	0.225	0.98	3.89	0.614
8 by 12	1	6	0.98	5.88	0.981	1.01	5.97	0.262
8 by 12	1	8	0.98	8.02	1.010	0.87	7.88	0.414
8 by 8*	1	4	0.95	4.20	0.454	0.97	3.78	0.462
8 by 8*	1	6	1.02	5.82	0.877	1.10	6.01	0.814
8 by 8*	1	8	0.98	8.32	1.030	0.98	8.08	0.656

a linear one in these cases. This was also true for a polynomial mean function whereby a quadratic was no longer significant and thus adjustment to avoid over fitting was unavoidable.

3. Except in a few instances, the estimate of σ^2 was generally higher in data with a trend compared to that without a trend. Fitting only covariance structure to data simulated with only trend (i.e. no incorporation of any covariance structure) was found to account for the trend perfectly well and there was no need for fitting a polynomial mean function or a smoothing spline. However, fitting a polynomial

Table 4.3: REML Covariance parameters for exponential errors

Array	Theoretical values		Estimated values					
	Parameter		*Mean (no trend)*			*Mean (+ trend)*		
	σ^2	a	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)
4 by 6	1	2	0.93	2.03	0.23	1.01	2.16	0.247
4 by 6	1	4	0.99	3.38	0.533	0.96	4.23	0.554
4 by 6	1	6	0.98	5.63	0.652	1.01	5.46	0.781
4 by 8	1	2	1.11	2.08	0.451	1.20	1.99	0.469
4 by 8	1	4	0.99	3.79	0.597	0.95	3.84	0.694
4 by 8	1	6	1.12	5.32	0.357	1.01	5.61	0.421
8 by 12	1	4	0.98	3.74	0.540	0.90	3.81	0.615
8 by 12	1	6	0.89	5.60	0.381	0.89	5.61	0.564
8 by 8*	1	4	0.95	3.74	0.816	0.89	3.88	0.665
8 by 8*	1	6	0.91	6.38	1.52	0.96	5.91	1.56

mean function or a smoothing spline alone seem to reduce the variance of contrasts i.e. it accounts for the trend better than covariance structure modelling.

General comments

The biases in estimating the value of ρ in AR(1) were less than 5% for both $\rho = 0.3$ and $\rho = 0.6$, while for AR(1)×AR(1) in some cases the biases were more than 10%. Generally, the same conclusions as those given above apply to AR(1) structure too. This confirms the theoretical basis of the models. Generally, convergence problems were less than 5%. The exponential error modelling agrees with AR(1) error modelling and either can be used to model the other. This is exactly the relation that was established by Gilmour *et. al.* (1997).

Table 4.4: REML Covariance parameters for Gaussian errors

Array	Theoretical values		Estimated values					
	Parameter		*Mean (no trend)*			*Mean (+ trend)*		
	σ^2	a	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)
4 by 6	1	2	0.98	2.04	0.321	0.89	2.40	0.641
4 by 6	1	4	0.90	3.45	0.337	0.78	3.37	0.278
4 by 6	1	6	1.24	6.41	0.632	1.26	6.36	0.781
4 by 8	1	2	0.92	1.84	0.864	1.06	1.83	0.642
4 by 8	1	4	0.97	3.32	0.189	1.01	3.38	0.144
4 by 8	1	6	0.94	6.50	0.860	0.98	6.53	1.431
8 by 12	1	4	0.91	3.74	0.447	0.97	3.74	0.421
8 by 12	1	6	0.86	5.64	0.911	1.04	6.02	1.34
8 by 8	1	4	0.95	4.20	0.454	0.97	3.68	0.462
8 by 8	1	6	0.89	5.66	1.34	1.06	5.89	1.32

It was observed that the variances of contrast or mean treatment differences were not significantly different if data simulated with a spherical model are fitted using ARMA(1,1) model in SAS PROC MIXED. However, no further investigation into the practical equivalence of ARMA(1,1) and spherical model was pursued in this work.

It was observed that, when a polynomial mean function is used, ρ or a is estimated accurately especially when the correct trend model or higher order trend model is fitted. However, if the true trend is under-fitted ρ or a is overestimated. There is thus some 'soaking up' of the spatial trend by the modelling of the error structure.

The above results about the validation of the suggested models for modelling systematic design data are reassuring. These methods are reliable and efficient in identifying

Table 4.5: REML Covariance parameters for AR and ARMA errors

Array	Theoretical values		Estimated values					
	Parameter		*Mean (no trend)*			*Mean (+ trend)*		
	σ^2	ρ/γ	$\hat{\sigma}^2$	$\hat{\rho}/\hat{\gamma}$	SD($\hat{\cdot}$)	$\hat{\sigma}^2$	$\hat{\rho}/\hat{\gamma}$	SD($\hat{\cdot}$)
AR(1)	1	$\rho = 0.6$	0.987	0.597	0.103	1.010	0.588	0.111
	1	$\rho = 0.3$	0.895	0.301	0.253	0.941	0.369	0.203
AR(1)	1	$\rho = 0.6$	0.946	0.521	0.301	0.091	0.507	0.142
×		$\rho = 0.6$		0.498	0.231		0.429	0.530
AR(1)	1	$\rho = 0.3$	0.753	0.273	0.040	1.303	0.259	0.154
		$\rho = 0.3$		0.238	0.054		0.331	0.112
ARMA	1	$\rho = 0.6$	0.879	0.594	0.201	0.990	0.542	0.142
(1,1)		$\gamma = 0.6$		0.643	0.231		0.629	0.330
	1	$\rho = 0.3$	1.030	0.333	0.120	1.303	0.340	0.114
		$\gamma = 0.8$		0.728	0.254		0.691	0.212

and modelling spatial variability. It is an indication that data from systematic designs that is assumed to contain spatial variability can easily be analysed and that the experimenter can reliably arrive at the same conclusions one would do if data was from a RCBD (without spatial variability). The importance of modelling trend and correlated errors is clearly underscored.

4.4 Beta-hat models

4.4.1 Introduction

Up to this point, discussion has centered around the assumption that the models are to be fitted to the whole experimental data. However these same methodologies can be applied to only subsets of data. For design II and design III in particular, these models can be fitted to each main plot by taking subplot plant populations as covariates. For design I, a single crop at a time can be analyzed. For example, to analyze the intercropping effect of crop 2 on crop 1, the the plant densities of crop 2 are used as covariates and models are fitted for each treatment combination of the spatial arrangement and the plant populations of crop 1. To evaluate the effect of crop 1 on crop 2, the reverse is taken. Since each main plot treatment/treatment combination is replicated, the fitted response curves/surfaces of different main plot treatments are then compared by use of parameter estimates. This is what is referred to as a *beta-hat model* (Milliken and Johnson, 2002). In beta-hat modelling two stage data analysis is done. The first stage involves obtaining of the regression parameters (beta's) and stage two involves modelling of the beta's. The idea of a beta-hat model is that in an experiment where there are more than one combinations of qualitative factors and many levels of a quantitative factor, observations from each combination of qualitative factors can be regressed on the levels of a quantitative factor independently. The different combinations of qualitative factors are then compared through modelling their respective regression parameters.

In the sections that follow, two beta-hat models termed model I and model II are introduced. Beta-hat model I allows further investigation of the effect of main plot treatments on the beta's (regression coefficients). Beta-hat model II is mainly for response surface exploration under different treatment combinations (e.g. main plot treatment combinations). This model helps in revealing where the maximum response occurs for

each treatment combination. The two models are complementary. The choice to use beta-hat models for modelling intercropping plant population experiments is due to their advantage in enabling efficient exploration of the response curve/surface. To motivate the discussion consider three factors a density levels of crop 1, b density levels of crop 2 and p levels of spatial arrangements.

4.4.2 Beta-hat Model I

Consider, for example, the effect of plant populations of crop 2 on crop 1 at a given level of spatial arrangement. The idea under beta-hat model I setting is to fit simple linear regression models to the subsets of data. Each subset of data $\mathbf{y}_{ikh}$ $(i = 0,\ldots,a,\ k = 0,\ldots,p;\ h = 1,2,\ldots,r)$ consisting of observations from ith density level of crop 1 and kth spatial arrangement in hth replicate and use b levels of crop 2 as variates/explanatory variables. Thus

$$\mathbf{y}_{ikh} = \mu_{ikh} + \beta_{ikh}X_j + \varepsilon_{ikh} \qquad \varepsilon_{ikh} \sim \text{iid}(0,\ \sigma^2) \tag{4.25}$$

where $j = 0,\ldots,b$ and X_j are the levels of crop 2 that constitute an explanatory variate. $\mathbf{y}_{ikh}$ is $b \times 1$ vector of observations for the (i,k,h)th combination at all levels of crop 2/explanatory variable. This also can be expressed in the simple form

$$\mathbf{y}_{ikh} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_b \end{bmatrix} \begin{bmatrix} \mu_{ikh} \\ \beta_{ikh} \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1kh} \\ \varepsilon_{i2kh} \\ \vdots \\ \varepsilon_{ibkh} \end{bmatrix} \tag{4.26}$$

With this setting therefore, a beta-hat model can be used to investigate the effect of spatial arrangement and crop *1* densities on the slope or additive effects, β_{ikh}, associated with factor combination (i,k,h), and the intercept μ_{ikh} using ANOVA model. The beta-hat model for the slopes can be expressed as

$$\hat{\beta}_{ikh} = \mu + \tau_k + \rho_i + (\tau\rho)_{ik} + \varepsilon^*_{ikh} \tag{4.27}$$

where μ, τ_k, ρ_i, and, $(\tau\rho)_{ik}$ denote overall mean, effect of spatial arrangement, effect of crop 1 density and their interactions respectively and ε^*_{ikh} is an error term. In the analysis of variance each parameter $\hat{\beta}_{ikh}$ is weighted using the square of its associated standard error. Let the inverse of this squared standard error be denoted as ω where $\text{Var}(\hat{\beta}_{ikh}) = \sigma^2\omega_{ikh}$. Thus $\varepsilon^*_{ikh} \sim \text{iid}(0, \sigma^2\omega_{ikh})$. A similar formulation can be constructed for the intercepts.

Estimation and test of hypotheses

Assuming treatment combinations are independent and assuming $\hat{\beta}_{ikh}$ are normally distributed, a two-way analysis of variance can be carried out using weighted least squares. The assumption of independence holds since within each main plot, treatments are allocated independently i.e. the choice of directions of systematic arrangement are independent. Since each parameter $\hat{\beta}_{ikh}$ is weighted using the square of its associated standard error model (4.27) can also be regarded as a *weighted linear* model. Since, the weights are inversely related to the variance, $\hat{\sigma}^2_{ikh}$, they reflect the information contained in $\hat{\beta}_{ikh}$. The smaller the variance $\hat{\sigma}^2_{ikh}$, the more the information $\hat{\beta}_{ikh}$ provides about the expected value of the slope. Denote a matrix W as a diagonal matrix of weights, i.e.

$$W = \begin{pmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_b \end{pmatrix}$$

and express (4.27) in matrix form as,

$$\boldsymbol{\beta} = F\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{4.28}$$

where F is a design matrix for $\boldsymbol{\theta}$ the vector of μ, τ_k, ρ_i and $(\tau\rho)_{ik}$ effects and $\boldsymbol{\varepsilon}$ is the error vector with variance-covariance matrix $\sigma^2 W$. Consider a positive definite matrix

K such that $K'K = W$ which leads to rewriting (4.28) as

$$K^{-1}\boldsymbol{\beta} = K^{-1}F\boldsymbol{\theta} + K^{-1}\boldsymbol{\varepsilon} \tag{4.29}$$

where $E(K^{-1}\boldsymbol{\varepsilon})= 0$ and

$$\begin{aligned} \text{Var}(K^{-1}\boldsymbol{\varepsilon}) &= K^{-1}\text{Var}(\boldsymbol{\varepsilon})K^{-1} \\ &= K^{-1}\sigma^2 W K^{-1} \\ &= \sigma^2 K^{-1}KKK^{-1} \\ &= \sigma^2 I \end{aligned}$$

Thus this set up satisfies the requirements of a linear model and ordinary least squares method (OLS) can be applied. The normal equations for weighted least squares or maximum likelihood estimation are given as

$$\begin{aligned} \left(F'K^{-1}K^{-1}F\right)\hat{\boldsymbol{\theta}} &= F'K^{-1}K^{-1}\boldsymbol{\beta} \\ \text{or } \left(F'WF\right)\hat{\boldsymbol{\theta}} &= F'W\boldsymbol{\beta} \end{aligned}$$

The estimators are obtained as;

$$\hat{\boldsymbol{\theta}} = (F'WF)^{-1}F'W\boldsymbol{\beta} \quad \text{and } \text{Var}(\hat{\boldsymbol{\theta}}) = (F'WF)^{-1} \tag{4.30}$$

Although in the above set up matrix $F'WF$ is assumed to be of full, it is also possible in some situation that matrix $F'WF$ is not of full rank. In such a situation only an estimable linear function of $\boldsymbol{\theta}$, say $\mathbf{L}\boldsymbol{\theta}$ is obtained where $\mathbf{L}' = \mathbf{a}'F$ and rank($\mathbf{L}$)$\subseteq$ rank($\boldsymbol{\theta}$). The normal equations would be given as

$$\boldsymbol{\theta}^0 = S^{g_1}F'W\boldsymbol{\beta} + (I - S^{g_1}S)\mathbf{z}$$

where $\mathbf{z}$ is arbitrary and S denotes $F'WF$ matrix while (g_1) denotes the generalized inverse of S. These equations can then be solved using constraints.

An F-test, that tests the hypothesis of common intercept and parallelism in slopes is easily implemented. In this way, the main effects of spatial arrangements, crop 1 densities and their interactions can be compared. A significant F-test associated with τ_k or ρ_i means different spatial arrangements or plant densities perform differently in presence of crop 2 plant densities. Of much interest, however, is the interaction of spatial arrangements and plant densities, $(\tau\rho)_{ik}$, that measures performance of different combinations of spatial arrangements and plant densities. The difference between response means from different combinations, $(\tau\rho)_{ik}$, are conducted pairwise using t-tests, least significant differences, etc. Optimum values/responses from different treatment combinations can be compared using t-test.

Incorporating spatial variability

Global trend

To account for possible trend in the field an appropriate function $f(t_{ih})$, which may be a polynomial mean function or a cubic smoothing spline in plot positions, t_{ih}, is included. That is, (4.26) becomes,

$$\mathbf{y}_{ikh} = \mu_{ikh} + \beta_{ikh}X_j + f(t_{ikh}) + \varepsilon_{ikh} \tag{4.31}$$

Consider modelling t_{ih} using a polynomial mean function. Then for model (4.31), the variance of $\mathbf{y}_{ikh}$ will be made up of two components i.e.

$$\text{Var}(\mathbf{y}_{ikh}) \;=\; \text{MSTrend} + \hat{\sigma}^2_{\varepsilon_{ikh}} \;=\; \hat{\sigma}^2_h + \hat{\sigma}^2_{\varepsilon_{ikh}}$$

The variance of $\hat{\beta}_{ikh}$, σ^2_{ikh}, is thus given by

$$\frac{(\hat{\sigma}^2_h + \hat{\sigma}^2_{\varepsilon_{ikh}})}{\sum(X_j - \bar{X}_.)^2}$$

In the case of fitting a smoothing spline for t_{ih} i.e. a semi-parametric model, the estimates for β_i and λ are obtained by minimizing a penalized quasi-likelihood function

$$\left\{\frac{1}{b}(\mathbf{y}_i - X\boldsymbol{\beta_i} - f(t))^2 + \lambda \int [f^m(t)]^2\right\} \tag{4.32}$$

where $\boldsymbol{\beta_i}$ is a vector of μ_{ikh} and β_{ikh} and $\mathbf{y}_i$ denotes $\mathbf{y}_{ikh}$. Both PROC GAM in SAS and gam in Splus fits this model and both μ_{ikh} and β_{ikh} are obtained with their respective standard errors. The selection of λ is done using Generalized cross validation (GCV) and the deviance is provided testing for adequacy. The fitting of only the parametric part of this model using GAM give the same results as those obtained using PROC GLM a standard linear regression procedure in SAS.

Correlated errors

To incorporate the spatial correlation in errors, the proposed error structure is AR(1), where

$$Cov(\varepsilon_{ijkh}, \varepsilon_{ij'kh}) = \sigma^2\rho^{j-j'} \quad \text{for } j \neq j'$$

Let $\boldsymbol{\varepsilon_i}$ be a vector of ε_{ikh} then (4.25) becomes

$$\mathbf{y}_i = X\boldsymbol{\beta}_i + \varepsilon_i \tag{4.33}$$

where $\varepsilon_i \sim N(0, \sigma^2 R(\boldsymbol{\theta}))$, and $R(\boldsymbol{\theta})$ is the variance covariance matrix that depends on spatial parameters $\boldsymbol{\theta}$ (ρ, h). The estimates of $\boldsymbol{\theta}$ are obtained using REML by maximising the residual likelihood

$$\ell(\boldsymbol{\beta}_i, \boldsymbol{\theta}) = -\frac{1}{2}\{ln|V| - (\mathbf{y}_i - X\boldsymbol{\beta}_i)'V^{-1}(\mathbf{y}_i - X\boldsymbol{\beta}_i) - ln|X'VX|\}$$

where $R(\boldsymbol{\theta})$ is denoted by V. The estimates of $\boldsymbol{\beta}_i$ are obtained using GLS as follows

$$\hat{\boldsymbol{\beta}}_i = [X'R(\boldsymbol{\theta})^{-1}X]^{-1}X'R(\boldsymbol{\theta})^{-1}\mathbf{y}_i$$

$$\hat{\sigma}^2 = \frac{(\mathbf{y}_i - X\boldsymbol{\beta}_i)'R(\boldsymbol{\theta})^{-1}(\mathbf{y}_i - X\boldsymbol{\beta}_i)}{(n-1)}$$

Therefore

$$\hat{\boldsymbol{\beta}}_i \sim N(\boldsymbol{\beta}_i, \sigma^2[X'R(\boldsymbol{\theta})^{-1}X]^{-1})$$

Another possible model for handling spatial variability is the ARMA(1,1) process whose semi-variogram is

$$\sigma^2(1 - \gamma\rho^{h-1})$$

where γ is a moving average parameter that smooths out spatial fluctuations in the observations. PROC MIXED in SAS fits both AR(1) and ARMA(1,1) error structures. The covariance model ARMA(1,1) was found to give closely similar mean estimates of treatments and treatment contrasts as models involving spherical and Gaussian models for data simulated with spherical and Gaussian error structures, respectively. Harris and Dallas (1996) through simulation studies indicated that $\hat{\boldsymbol{\beta}}_i$ is less biased even when $R(\boldsymbol{\theta})$ is not known. These comments give confidence about the use of the proposed models.

Joint Error + Trend model

The only covariance model that can easily be applied in such a situation is the ARMA (1,1). This model was found to perform quite well through simulation studies.

Notice that in incorporating spatial variability in beta-hat model I only AR and ARMA covariance models can easily be fitted since the other models such as spherical, exponential and Gaussian require specification of a grid. The analysis under bet-hat model I set up assumes only a single row/column layout. In other words the data structure has no grid set up.

After accounting for spatial variability, the parameter estimates are then used to construct model (4.27) or (4.28). Further analysis is as indicated under this model.

4.4.3 Beta-hat Model II

Note that in model set up I, the regression model (4.25) is fitted to each factor combination responses from each replicate separately. Now consider fitting the same simple regression model for each factor combination but over all its replicates. That is consider the model

$$\mathbf{y}_{ik} = \mu_{ik} + \beta_{ik} X_j + \varepsilon_{ik} \qquad \varepsilon_{ik} \sim \text{iid}(0,\ \sigma^2) \tag{4.34}$$

where $\mathbf{y}_{ik}$ is $br \times 1$ vector of observations from ith level of crop 1 density and kth spatial arrangement. Clearly, the μ_{ik} and β_{ik} are functions of treatment combinations. This model set up is a fit of separate regression or response lines/curves to different treatment combinations i.e response curve fitting. Assuming independence of parameters from each response curve, formal t-tests can be conducted to test for separate curves or parallelism in slopes i.e. to test whether the slopes or intercepts are significantly different from each other. Response curve fitting has an advantage of revealing the pattern of the response as plant populations increase.

The t-test is conducted on the hypothesis;

$$H_0 : \beta_{11} = \beta_{12} = \beta_{13} = \ldots = \beta_{ap} = \beta_n \tag{4.35}$$

Logically, this hypothesis is rejected if there exists at least one pair $(i\,j)$ such that

$$H_A : \beta_{ik} \neq \beta_{i'k'} \quad i \neq i',\ k \neq k'$$

Formally, the appropriate t-test of difference between $\hat{\beta}_{ik}$ and $\hat{\beta}_{i'k'}$ is given by

$$\frac{\hat{\beta}_{ik} - \hat{\beta}_{i'k'}}{\sqrt{\text{Var}\left(\hat{\beta}_{ik} - \hat{\beta}_{i'k'}\right)}}$$

where $\text{Var}\left(\hat{\beta}_{ik} - \hat{\beta}_{i'k'}\right) = \text{Var}(\hat{\beta}_{ik}) + \text{Var}(\hat{\beta}_{i'k'})$ since $\text{Cov}\left(\hat{\beta}_{ik}, \hat{\beta}_{i'k'}\right)$ is assumed to be zero.

Weighted regression (F-test) on the β_{ik}'s can also be conducted to test for differences between them. Similarly, the same weighted regression using standard errors as weights

can be done on μ_{ik}'s. The significant differences in μ_{ik}'s using F-test or t-test indicates the difference in the performance of main plot treatments in absence of interaction with 'systematic' subplot treatments.

Incorporating spatial variability in beta-hat model II takes the same steps used as in general spatial models in the previous section. Models (4.3), (4.11), (4.16), (4.22) or (4.23) are used to model spatial variability with the only exception that fixed effects matrix, X, contains only systematically arranged crop densities (subplot treatments). The treatment effect vector, $\boldsymbol{\beta}$, contains only the effects of subplot treatments. Thus, regression of crop responses on subplot treatments is implemented with spatial variability incorporated and the generated beta's are compared using F or t-test.

4.4.4 Beta-hat model validation

The data simulated in the previous section were used in validation. The data were fitted using beta hat model I and II. The results are presented in Tables 4.6 and 4.7. For model set up I, the ARMA(1,1) model was used to fit data with spherical error structure while the AR(1) model was also used to fit data with exponential error structure.

The results presented are reassuring in that beta hat models can be fitted and appropriate covariance models are still estimable. The beta-hat models are as good as general models. Results from Table 4.6 and Table 4.7 indicated the following observations.

1. In beta-hat model I, the underestimation of σ^2 at low values of ρ (i.e. $\rho < 0.3$) was observed in 21% of the total of 600 samples. The use of this model needs care and a pragmatic data exploration approach is necessary. A general pragmatic approach is included in next section.

2. In beta hat model II, the results for exponential and Gaussian model were similar to those of the spherical model, that is, similar trend as under general models. No-

Table 4.6: REML Covariance parameters for AR and ARMA errors

Array	Theoretical values		Estimated values					
	Parameter		*Mean (no trend)*			*Mean (+ trend)*		
	σ^2	ρ	$\hat{\sigma}^2$	$\hat{\rho}$	SD($\hat{\rho}$)	$\hat{\sigma}^2$	$\hat{\rho}$	SD($\hat{\rho}$)
AR(1)	1	0.6	1.210	0.586	0.091	1.070	0.580	0.079
	1	0.4	1.010	0.396	0.014	1.000	0.386	0.061
	1	0.3	0.675	0.327	0.123	0.864	0.369	0.207
ARMA	1	0.6	1.230	0.591	0.201	0.990	0.598	0.042
(1,1)	1	0.3	0.970	0.321	0.025	0.980	0.309	0.030
	1	0.8*	1.230	0.833	0.091	1.210	0.940	0.204
	1	0.6*	0.970	0.528	0.284	0.870	0.691	0.438

tice that the covariance parameter estimates are not as precise as those estimated with general models.

From the results above, beta-hat models are shown to perform like general spatial linear mixed models in that they effectively extract spatial variability. Their use, therefore is reliable and they are also easily implemented. These results underscore the importance of beta-hat models in handling systematic design data from intercropping experiments involving plant populations. beta-hat model II in particular has an advantage of enabling effective response surface/pattern exploration.

Table 4.7: Covariance parameters for Spherical errors for Model II

Array	Theoretical values		Estimated values					
	Parameter		*Mean (no trend)*			*Mean (+ trend)*		
	σ^2	a	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)	$\hat{\sigma}^2$	$\hat{a}$	SD($\hat{a}$)
4 by 6	1	2	0.867	1.911	0.440	1.054	1.854	0.392
4 by 6	1	4	1.240	3.948	0.167	1.186	3.692	1.082
4 by 6	1	6	1.259	5.798	1.036	1.306	5.650	1.197
4 by 8	1	2	1.024	5.728	1.024	1.218	5.871	0.995
4 by 8	1	4	1.129	3.948	0.987	1.082	3.538	0.670
4 by 8	1	6	1.116	5.725	1.135	1.142	5.692	1.122
8 by 12	1	4	0.978	4.289	0.813	1.369	4.342	0.941
8 by 12	1	6	1.155	5.710	1.120	1.163	5.858	1.119
8 by 12	1	8	1.280	7.708	0.243	1.285	7.768	1.013

4.5 Data modelling: a pragmatic approach

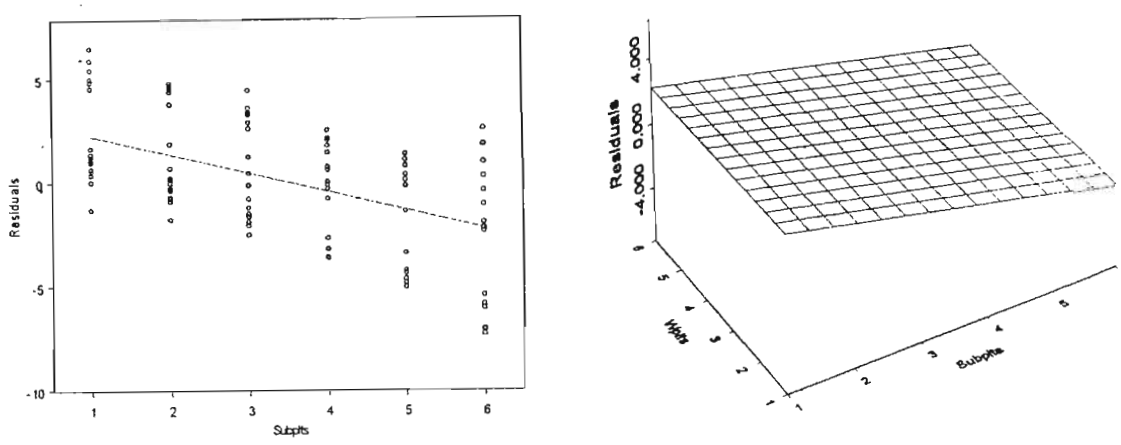
In this section three simulated data sets will be used to illustrate the steps that should be involved in model building. These are Monte Carlo realizations with the following properties;

1. Data S1: Monte Carlo sample with a global trend in subplots $(0.9+0.7R_i-0.3R_i^2)$ for a 4×8 array (global-trend-only data).

2. Data S2: Same Monte Carlo sample as in S1 with a global trend excluded. Instead spherical random process errors $(a = 4, \sigma^2 = 1)$ were included to give correlated observations (correlated-errors-only data).

3. Data S3: A combination of global trend and spherical error process were simulated.

When building up a model, because the decomposition of the observed process is not unique, the first step is to look for large scale variation or trend in the data. One of the methods is to fit a model to the data using OLS to estimate the residuals. The second step involves computing residuals, $\xi^*_{s_i} = \mathbf{y}_i - \hat{\mathbf{y}}_{(s_i)}$, for $i = 1, \ldots, n$, and plotting them against the subplots or columns and rows. Figure 4.1(a) and (b) are plots of residuals from S1. Clearly, there is a downwards trend i.e. decreasing trend with subplot number. Also from Figure 4.1(b), it can be observed that the there is no trend in rows/mainplots. After identifying such a trend residuals can be treated as response variates to the rows and columns and using OLS method an appropriate polynomial mean function can be determined. Determination of an appropriate final mean function involves two steps. Fit the residuals as indicated above and then include the function in data modelling. Then delete the higher order terms in rows or subplots that are not significant according to the F or t-test. This will lead to obtaining a parsimonious model.



(a) (b)

Figure 4.1: Plot of OLS residuals for data S1

Another way of investigating a trend is by simple plots (trellis) whereby the yields are plotted against rows and columns per replicate. Any difference in the shape of the plots may provide an insight into the trend in the data. Figure 4.2 for S1 is an example of a plot of yields against subplots per replicate. The block effects can clearly be observed. Notice immediately that the shape of distribution of yields in the third replicate (Rep 3) is different from those in Rep 1.

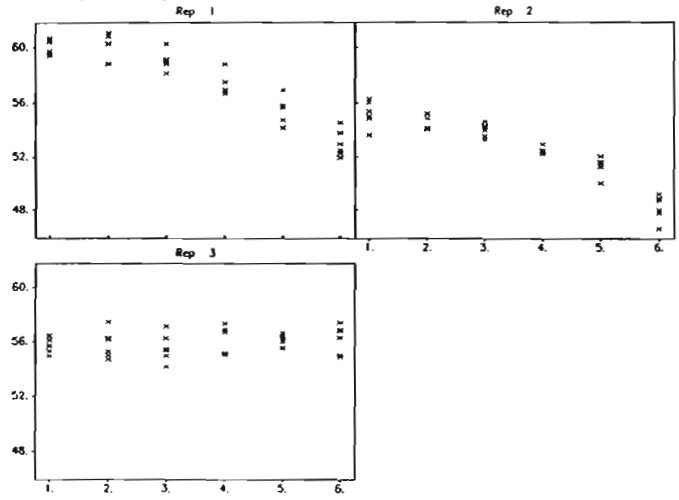


Figure 4.2: Trellis plots for data S1

The next step is to explore for spatial dependence between the observations. The residual-based variogram or covariogram estimates provide an important tool in identifying the spatial dependence. For example, under an isotropic process with $E[\xi^*(s_i)] = 0$, the variogram can be computed as

$$\hat{\gamma}_{\xi^*}(h) = \frac{1}{|2N(h)|}\sum^{N(h)}\left\{\hat{\xi}(s_i) - \hat{\xi}(s_j)\right\}^2, \qquad h = s_i - s_j$$

where $N(h)$ is the number of distinct pairs of plots separated by h and $\xi^*(s_i)$ are residuals at position s_i. This step is important as it gives an indication about possible spatial covariance structure that may be appropriate. Consider data S2 whose variogram is given in Figure 4.3(a). If there was no spatial dependence in the data the variogram would remain constant as lag increases. Thus modelling will have to take into account this spatial dependence. The shape of the empirical or sample variogram gives an idea about selection of an appropriate covariance models. Non-linear least squares or REML can be used to fit different covariance models and selection of the best can be done. In the present example REML is used to fit the covariance models.

Furthermore, the behavior of the residual based variogram or covariogram may reveal non-stationarity of the residual random process. The presence of a trend for example may result in a quadratic increase of the variogram estimator in accordance with the lag. Consider for instance, Figure 4.3(b) for S3. The variogram seem to increase indefinitely with increasing lag. In such a case a trend model have to be adopted. Note that if the trend has been successfully identified in the first step (i.e. using trellis or residual plots), the variogram or covariogram would be based on residuals after adjusting for the trend.

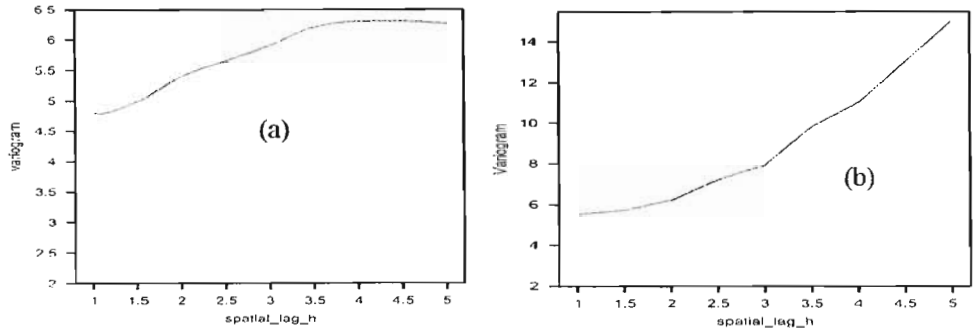


Figure 4.3: Sample variograms for data S1 and S3

The final step is model checking for the adequacy of the covariance model and the nature of the trend fitted. This is done using such tools as LRT, AIC, BIC and residual plots. After exploring data for global trend, packages such as SAS PROC MIXED can also enable one to determine the presence of spatial dependence in the data. Fit the data without any covariance model and then fit the data with various covariance models. The changes in LRT or deviance, AIC and BIC statistics can be used as an indicators of adequacy. An improvement in this statistics on fitting covariance models will imply a possible presence of spatial dependence in the data. Variogram exploration can also be done. Consider fitting the true covariance model (spherical) and alternative models for data S2. Clearly from the Table 4.8, the spherical model performs well according to AIC and BIC criteria (have the highest AIC and BIC values). Notice that the fit statistics of ARMA(1, 1) are close to those of spherical covariance model. In this data, ARMA(1, 1) is the second best model according to AIC, AICC and BIC criteria.

Table 4.8: REML covariance parameter estimates for data S2

Error	Cov. Parameters		Fit statistics			
model	$\hat{\sigma}^2$	$\hat{a}$	LRT	AIC	AICC	BIC
Exponential	0.7127	1.30	214.40	247.40	247.70	244.60
Gaussian	0.6723	1.01	242.90	248.90	249.20	246.20
Spherical	0.9645	4.02	**244.10**	**250.10**	**250.40**	**247.40**
AR(1)	0.7127	$\rho = 0.46$	214.40	247.40	247.70	244.60
ARMA(1,1)	0.9598	$\rho = 0.42$	241.30	249.3	249.8	245.7
		$\gamma = 0.46$				

4.6 Case studies

To illustrate the previous sections, a data set from an experiment conducted by students at Makerere University in Uganda is considered. The cowpea - sorghum intercrop experiment was set up in a parallel-row design. Two spatial arrangements (SA), two crop density levels of cowpea (dC) and eight sorghum crop density levels (dS) were involved in the experiment. The sorghum density levels were systematically arranged in subplots while combinations of cowpea densities and spatial arrangements were randomly allocated to the main plots. The interest was to determine intercropping advantages through examining the response pattern. Only dry matter residuals were available for analysis. Recall that the analysis of relative LER, RE and total yield are handled in the same way. For example, relative LER or RE will entail combining the two crop yields together and the variates so formed are modelled in a similar way like the yield (see Section 2.2). The data is included in Appendix D.

Data exploration

In the first step, the large scale variation or trend was considered for the data. SAS PROC MIXED is used to fit the data with treatment effects only and the residuals are obtained. Residuals were then plotted against rows (main plots) and columns (subplots) (Figure 4.4). No trend was observed in residuals. The histograms of the residuals do not show any outlier (Figure 4.5; (a) for cowpea yield and (b) for sorghum yield). Figure 4.6(a) and (b) are two dimensional grayscale plots of residuals in field positions for cowpea and sorghum, respectively. There is no clear indication of any drift or trend effect. The plot looks fairly uniform. This can be confirmed by trellis plots (Figure 4.7; (a) for cowpea and (b) for sorghum) where all the blocks have the same shape of yield distribution.

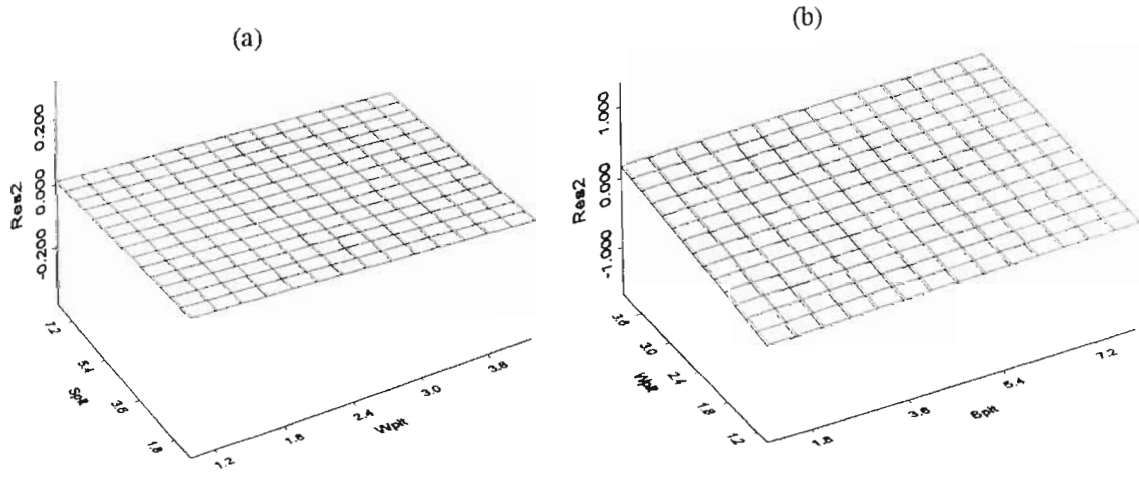


Figure 4.4: Plot of residuals against main plots and subplots

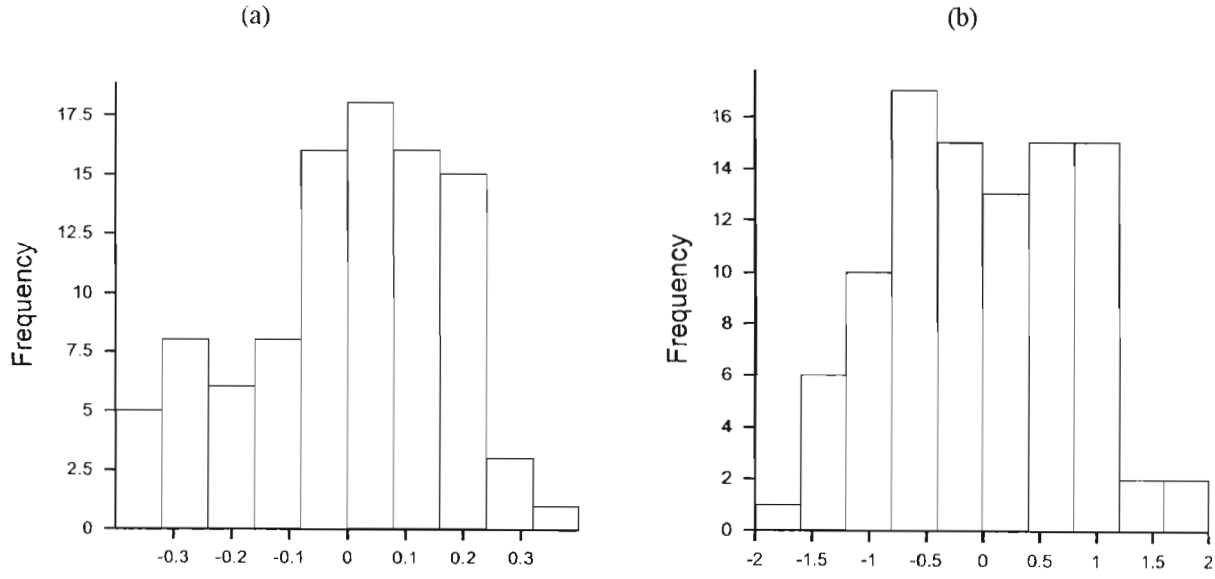
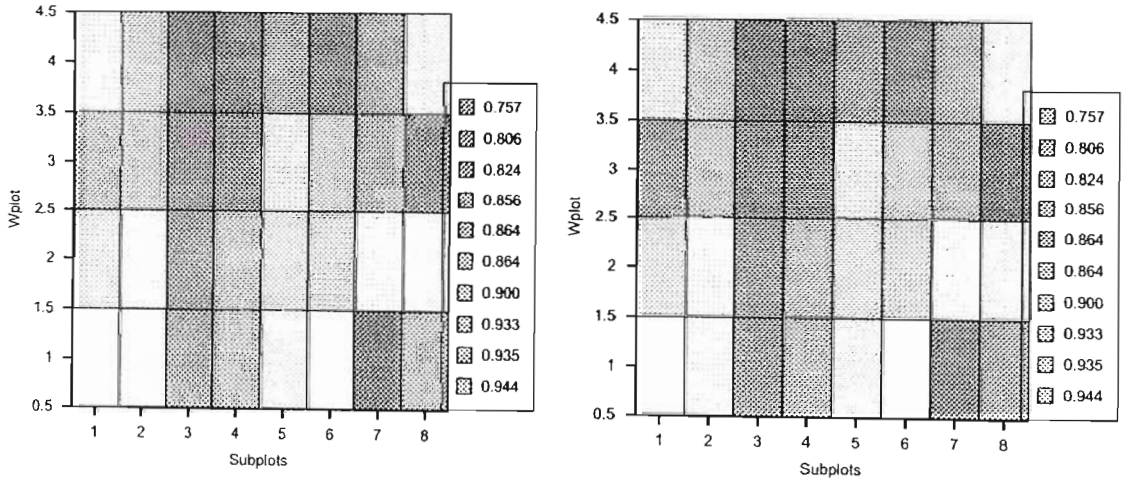


Figure 4.5: Histogram of residuals for cowpea-sorghum intercrop



(a) Cowpea yields (b) Sorghum yields

Figure 4.6: The gray scale for cowpea-sorghum intercrop yields

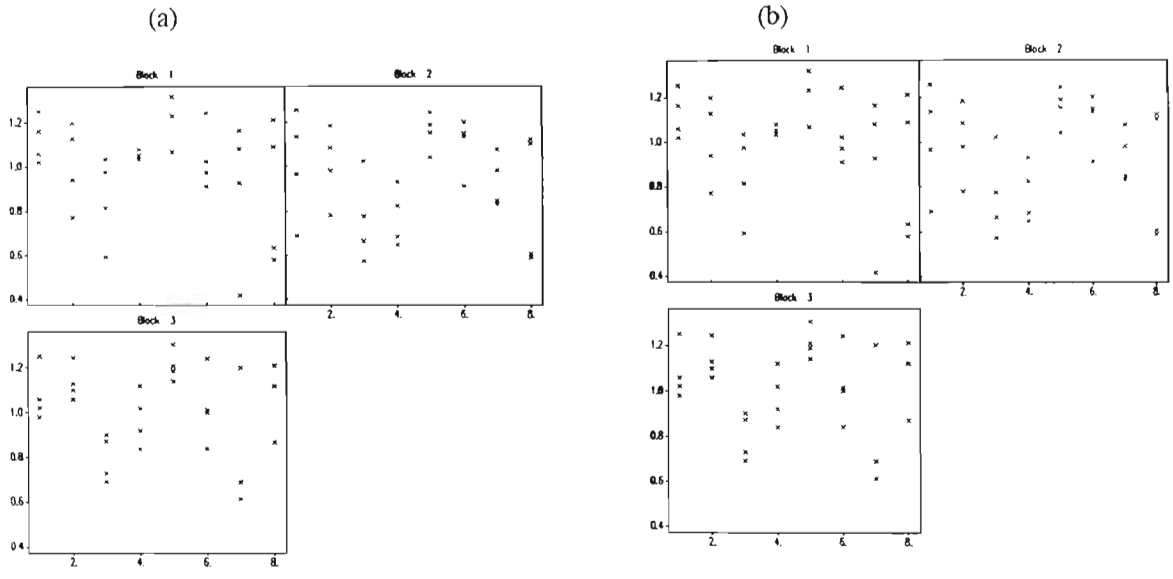


Figure 4.7: The trellis plots for cowpea-sorghum intercrop yields

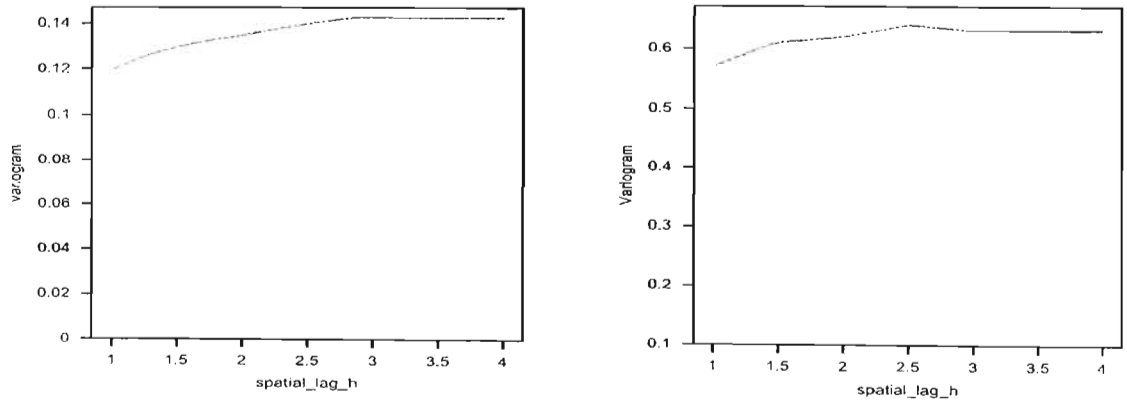


Figure 4.8: Sample variograms for cowpea-sorghum intercrop yields

Figure 4.8(a) and (b) are sample variograms of the cowpea and sorghum data, respectively. The increase in the variograms before levelling off is an indication that there is spatial dependence in the data sets. These plots also confirms that there is no trend effect. From this figure, the sorghum data seem to have a slightly high sill than the cowpea data. The sills for the two data sets are expected to be the same since the data are taken on the same plots. However, without loss of generality, these sills can be taken to be the same.

Data analysis: Spatial linear mixed model

Given that sorghum and cowpea data have the same properties, only one data set (cowpea yield data) anlaysis was reported here. The advantage of similar properties of crop yield observations is that the two crop yields can be combined into relative LER or RE to obtain accurate results. The REML method was used to fit covariance models to the data. Model (4.3) was used to fit the data. The parameter estimates and fit statistics from PROC MIXED are presented in Table 4.9. The nugget effect model alone and nugget effect and covariance models were fitted. Only covariance models without a nugget variance are presented. This is because all the covariance models with nugget variance had lower AIC and BIC values compared to models without the nugget variance. The value of the LRT is higher in a model with nugget effect only. This underscores the necessity of modelling spatial dependence between observations. The AIC, BIC and LRT suggest the ARMA(1, 1) model as adequate. Table 4.10 provides the statistical analysis of cowpea data using PROC MIXED. All fixed effects are not significant at $\alpha = 0.05$.

Table 4.9: REML covariance parameter estimates for cowpea data

Error	Cov. Parameters		Fit statistics			
model	$\hat{\sigma}^2$	$\hat{a}$	LRT	AIC	AICC	BIC
Nugget	0.3250	-	191.30	147.30	147.60	144.60
Exponential	0.3893	2.34	150.90	146.90	147.20	144.20
Gaussian	0.3286	1.16	160.70	156.70	157.00	154.00
Spherical	0.5132	5.46	155.9	151.90	152.20	149.20
AR(1)	0.3893	$\rho = 0.65$	150.90	146.90	147.20	144.20
ARMA(1, 1)	0.4199	$\rho = 0.87$	148.20	**156.20**	**156.30**	**152.60**
		$\gamma = 0.67$				

Table 4.10: PROC MIXED type III tests of fixed effects

Effect	Num df	Den df	F value	Pr > F
dC	1	84	0.13	0.7214
dS	1	2	0.06	0.8236
dC*dS	1	84	0.34	0.5639
SA	1	84	0.31	0.5818
SA*dC	1	84	0.02	0.8871
SA*dS	1	84	0.34	0.5607
SA*dC*dS	1	84	0.09	0.7698

Data analysis: beta-hat models

Consider using beta-hat models. Fitting beta-hat model II and testing for appropriate covariance models on average AIC and BIC values led to the conclusion that ARMA(1,

Table 4.11: PROC MIXED t-tests of fixed effects for beta-hat model II

Treatment	$\hat{\beta}_0$	s.e.	$\hat{\beta}_1$	s.e.	t-test for $\hat{\beta}_0$	t-test for $\hat{\beta}_1$
1:1	4.4833	0.7495	0.000613	0.050440		
2:1	4.8075	0.4695	-0.029740	0.044550	0.7271	0.6491
1:2	4.3256	0.2996	0.054900	0.056410	0.8430	0.4851
2:2	4.6086	0.4698	0.003329	0.038900	0.8850	0.9425

1) is adequate. For example, ARMA(1, 1) had AIC and BIC values of 42.85 and 38.98, respectively followed by Gaussian model with 42.00 and 38.29, respectively. Fitting a simple linear regression and modelling spatial variability using ARMA(1, 1) led to the following results (Table 4.11) for response surface parameters. The tests are in reference to beta values at treatment combination level 1 : 1. All these tests and other comparisons were not significant.The same conclusions for SA*dS and SA*dC*dS are reached at as was in using general spatial linear mixed model.

Beta-hat model I was also fitted with spatial variability modelling through ARMA(1,1). Although, for some (2 of 12 subdata sets) subdata sets the final hessian matrix was not positive definite, the following results suffice to compare the fixed effects (Table 4.12). Again tests on effect of SA and dC on the intercepts (β_0's) and slopes (β_1's) shows no significance at $\alpha = 0.05$. The same conclusion arrived at using general spatial linear mixed models are also arrived at using beta-hat model I.

From the preceding sections, it has been shown that it is quite imperative that systematic design data starts with data exploration. The pragmatic approach illustrated above have been shown to be effective in identifying data properties. Modelling spatial correlation and global trend is an efficient method for data with spatial variability (especially data from the systematic designs). This avoids loss of information due to

Table 4.12: PROC MIXED fixed effect tests for beta-hat model I

			Test on intercepts		Test on slopes	
Effect	Num df	Den df	F-value	Pr > F	F-value	Pr > F
SA	1	6	0.710	0.4304	0.350	0.5768
dC	1	6	0.100	0.7601	3.420	0.1141
SA*dC	1	6	0.030	0.8747	1.110	0.3329

imprecisely estimating error variance and treatments. The models incorporating spatial variability in data analysis are easily implemented in available statistical software such as Genstat, SPLUS and SAS. Therefore, data with spatial variability is no longer a problem during analysis. In this sense, therefore, systematic designs that have many advantages in intercropping experiments involving plant populations can be used. This is because the only assumed disadvantage of spatial variability has been shown to be easily handled using the models suggested in this chapter.

Chapter 5

Evaluation of Design Efficiency

5.1 Introduction

The efficiency of systematic designs relative to randomised complete block design (RCBD) is considered. In addition to the evaluation of design efficiency the cost effectiveness of systematic designs, otherwise called efficiency under a cost function, is evaluated. Section 5.2 introduces general design efficiency theory, while in Section 5.3 specific efficiency measures are presented and simulation results reported. Section 5.4 deals with cost effectiveness of systematic designs in intercropping experiments.

5.2 Design efficiency theory

The efficiency of a design is its precision relative to an orthogonal design with the same number of treatment replications. It is a measure of how precisely treatment means or contrasts are estimated in a particular design. Efficiency is defined by the amount of information, which is the reciprocal of error variance, obtained in an experiment. The higher the information obtained from design ξ compared to the others from a set of competing designs χ, the higher the efficiency of ξ. Designs which minimize global

measures of variance on treatment contrasts called optimality criteria (Kiefer, 1959) are desired.

The classical formulation of optimal design theory is for a linear regression problem in which certain variables x_i are chosen by the experimenter and p functions of interest are known functions of x_i denoted $f_1(x_i), f_2(x_i), \ldots, f_p(x_i)$. The ith data point, y_i, is then

$$y_i = \sum_{j=1}^{p} f_j(x_i)\beta_j + \varepsilon_i \qquad i = 1, 2, \ldots, n$$

where $\beta_1, \beta_2, \ldots, \beta_p$ are unknown parameters and ε_i are the uncorrelated errors with mean 0 and common variance σ^2. Writing

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad F = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \ldots & f_p(x_1) \\ f_1(x_2) & f_2(x_2) & \ldots & f_p(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \ldots & f_p(x_n) \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \tag{5.1}$$

the best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$ is given by $(F'F)^{-1}F'\mathbf{y}$, with covariance matrix $(F'F)^{-1}\sigma^2$. If $\boldsymbol{\xi}$ is viewed as an arbitrary probability measure on the experimental space χ, which puts weight $\frac{1}{n}$ at each of the design points $x_1, \ldots, x_n$, then

$$\frac{1}{\sigma^2} F'F = \frac{1}{n}\sum_{i=1}^{n} f(x_i)f(x_i)' = \int_{\chi} f(\mathbf{x})f(\mathbf{x})'d\xi(\mathbf{x}) = \frac{1}{\sigma^2}M(\xi) \tag{5.2}$$

where $M(\xi)$ is the *information matrix* of $\boldsymbol{\beta}$ and $f(\mathbf{x}) = \big(f_1(x), f_2(x), \ldots, f_3(x)\big)'$. The integral in the above equation is called the Stieltjes integral. Matrix $\mathbf{M}(\xi)$ is a measure of accumulated 'precision' and for linear regression the variance-covariance matrix of the least squares estimates $\hat{\boldsymbol{\beta}}$ is the inverse of this matrix i.e. $\text{var}\{\hat{\boldsymbol{\beta}}\} = \frac{\sigma^2}{n}M(\xi)^{-1}$. The variance of the predicted response at x_i, is given as $d(\mathbf{x}, \xi) = f(\mathbf{x})'\mathbf{M}(\xi)^{-1}f(\mathbf{x})$. Thus

choosing a good design means making $M(\xi)$ 'large' in some suitably defined sense and this forms the basis of optimality criteria (Pukelsheim, 1993).

The typical information matrix functionals are the matrix means ϕ_p, for $p \in [-\infty; 1]$ called ϕ-efficiency measures.

The ϕ-efficiency of a design $\xi \in \chi$ is defined by

$$\phi_p(\xi) = \frac{\phi[C_c(M(\xi))]}{v(\phi)}$$

It is a number between 0 and 1, and gives the extend to which design ξ exhausts the maximum information $v(\phi)$ for $\boldsymbol{c'\beta}$ (c being a contrast matrix) with information matrix $C_c(M(\xi))$. The special forms of ϕ_p criteria include D-, A- and E-optimality for $p = 0, 1, \infty$ (Pukelsheim, 1993) which are defined below.

Definition 5.1 *D-Optimality (ϕ_0) is the determinant criterion defined as*

$$\min_{\xi}\{-log|M(\xi)|\}$$

i.e. the design that minimizes the generalized variance of the parameter estimators, or equivalently minimizing the volume of a confidence ellipsoid region. Equivalently a design that maximizes the product of eigenvalues of the matrix $C_c(M(\xi))$, $\Pi_i\lambda_i$ is D-optimal.

Definition 5.2 *A-Optimality (ϕ_1) is the trace criterion. It minimizes the average variance of the optimal parameter estimators. It is defined as max trace$(C_c(M(\xi)))$.*

Definition 5.3 *E-Optimality (ϕ_∞) is the smallest eigenvalue criterion: It minimizes the variance of the well estimated contrast or equivalently minimizes the max$(1/\lambda_i)$.* It minimises

$$\max_{c'c=1} \text{Var}\left(c'\hat{\theta}\right)$$

over the variances of all (normalised) linear combinations of parameter estimates $(c'\hat{\beta})$.

In a set of designs, χ, a design $\xi \in \chi$, with the lowest ϕ_p value is considered optimum and thus more efficient than the others.

In the case of correlated errors with covariance matrix V the variance-covariance matrix of $\hat{\beta}$ is given by

$$\mathrm{Var}(\hat{\beta}) = \left(FV^{-1}(\xi)F'\right)^{-1}$$

and thus the information matrix is defined as

$$M(\xi) = FV^{-1}(\xi)F' \tag{5.3}$$

This information matrix (5.3) is not a sum of information matrices of single observations as that under uncorrelated errors (5.2). It cannot be used directly in convex theory, which is essentially based on additivity of information matrices. Thus, the optimal designs cannot be solved analytically but can be determined numerically. Federov and Müller (1989) and Müller (2000) devised the use of an *approximate information matrix*, in algorithms for constructing spatial optimal designs, which is very close to the exact information matrix. In this work the efficiency of systematic designs is done numerically. In the next section a discussion of evaluation of systematic designs follows.

5.3 Estimation of design efficiency

In the present study the interest is in the evaluation of efficiency of systematic designs relative to RCBD based on optimality criteria reviewed in the previous section. Each systematic design will be assessed through a large number of data simulations and analysing the resultant data sets and comparing the results with data simulations for RCBD. Data analyses is to be done mainly in SAS and GenStat using PROC IML, PROC MIXED, PROC GLM, PROC GAM in SAS and REML, ANOVA in GenStat.

5.3.1 ϕ_p - directed efficiency

Suppose interest is in drawing inferences regarding L independent parametric functions of $\boldsymbol{\beta}$. The ϕ_p−directed relative efficiency of systematic designs to RCBD is then based on the information regarding $L\boldsymbol{\beta}$. The contrast matrix L consists of comparisons between main plot treatments, their interactions and also their interactions with split plot treatments. The justification for defining L as above is from the assumption that the researcher is not interested in testing for significance differences between split plot treatments (systematically arranged plant densities). Basing on models given in the previous chapter, the relative efficiencies are defined through the functionals of the information matrices. Note that under assumptions of normality the variance-covariance matrix of fixed effects estimates is the inverse of information matrix. That is if

$$C^{-1} = \mathrm{Var}(\boldsymbol{\beta}) = \left(XH^{-1}X\right)^{-1} \text{ then } M(\xi) = XH^{-1}X$$

Each of the models suggested in Chapter 4 has information matrix conditional on components $(\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\psi}')$ present in a model of the form

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

Let $C_d(\xi)$ and $C_R(\xi)$ denote information matrices for $\boldsymbol{\beta}$ conditional on covariance components $(\boldsymbol{\alpha}', \boldsymbol{\gamma}', \boldsymbol{\psi}')$ and $(\boldsymbol{\psi})$ conditional on $(\boldsymbol{\alpha}', \boldsymbol{\beta}', \boldsymbol{\gamma}')$, respectively where $C_d(\xi) = C_{[11.2]}$, $C_R(\xi) = C_{[22.1]}$ and $C_{[ii.j]} = C_{ii} - C_{ij}C_{jj}^{-}C_{ji}$, and where $(\cdot)^{-}$ denotes any generalised inverse (true inverse for nonsingular or pseudo-inverse for singular C_{jj}). From this setting the efficiency measures are defined using functions of $C_d(\xi)$ and $C_R(\xi)$, the treatment and residual effects information matrices. The information matrix of $\boldsymbol{L\beta}$, for example, would then be

$$\left\{L'(C_d)^{-}L\right\}^{+}$$

where $(\cdot)^+$ denotes the Moore-Penrose inverse (Azzalini and Giovagnoli, 1987). The matrix $(C_d)^-$ can be also be partitioned as

$$\begin{bmatrix} M_{11}(\xi) & M_{12}(\xi) \\ M_{21}(\xi) & M_{22}(\xi) \end{bmatrix}$$

Then the information matrix of $L\boldsymbol{\beta}$ is given by

$$M_L^*(\xi) = M_{11}(\xi) - M_{12}'(\xi)M_{22}^-(\xi)M_{21}(\xi) \tag{5.4}$$

Suppose RCBD is denoted as design $\xi_{\mathbb{R}}$ and a systematic design as design $\xi_{\mathbb{S}}$. Suppose also that the identified dispersion matrix, an inverse of a matrix $M_L^*(\xi)$ in (5.4) for design $\xi_{\mathbb{S}}$ is $C_L(\xi_{\mathbb{S}})$ from spatial modelling and that of design $\xi_{\mathbb{R}}$ is $C_L(\xi_{\mathbb{R}})$. Then the Fisher efficiency (considered for easy of computation) is given as

$$\mathrm{E}[L'\tilde{\boldsymbol{\beta}}(\xi_{\mathbb{S}}), L'\tilde{\boldsymbol{\beta}}(\xi_{\mathbb{R}})] \quad = \quad \frac{L'\big(C_L(\xi_{\mathbb{R}})\big)L}{L'\big(C_L(\xi_{\mathbb{S}})\big)L} \tag{5.5}$$

Also if the F-test is considered for $H_0 : L\boldsymbol{\beta} = \boldsymbol{\delta_0}$ against $H_{\mathrm{A}} : L\boldsymbol{\beta} \neq \boldsymbol{\delta_0}$ under designs $\xi_{\mathbb{S}}$ and $\xi_{\mathbb{R}}$, the Pitman efficiency of design $\xi_{\mathbb{R}}$ relative to $\xi_{\mathbb{S}}$ in testing H_0 is given as the ratio $E(\xi_{\mathbb{R}}, \xi_{\mathbb{S}}|L) = \lambda_L(\xi_{\mathbb{R}})/\lambda_L(\xi_{\mathbb{S}})$ i.e. a ratio defined by their non-centrality parameters (λ_L's) (Jensen and Ramirez, 1993). Thus

$$E(\xi_{\mathbb{R}}, \xi_{\mathbb{S}}|L) = \frac{(L\boldsymbol{\beta} - \boldsymbol{\delta_0})'\big\{L'\big(C_L(\xi_{\mathbb{S}})\big)L\big\}^{-1}(L\boldsymbol{\beta} - \boldsymbol{\delta_0})}{(L\boldsymbol{\beta} - \boldsymbol{\delta_0})'\big\{L'\big(C_L(\xi_{\mathbb{R}})\big)L\big\}^{-1}(L\boldsymbol{\beta} - \boldsymbol{\delta_0})} \tag{5.6}$$

with $H_0 : L\boldsymbol{\beta} - \boldsymbol{\delta_0} = c$, (5.6) assumes the form

$$E(\xi_{\mathbb{R}}, \xi_{\mathbb{S}}|L) = \frac{\mathbf{c}'\big\{L'\big(C_L(\xi_{\mathbb{S}})\big)L\big\}^{-1}\mathbf{c}}{\mathbf{c}'\big\{L'\big(C_L(\xi_{\mathbb{R}})\big)L\big\}^{-1}\mathbf{c}} \tag{5.7}$$

Based on (5.5) and (5.7), the average of variances of contrasts of interest and mean differences in the systematic design relative to the same contrasts in RCBD were used (*relative A−directed efficiency*). The largest variances of contrasts and mean differences

(*relative E−directed efficiency*) are also reported. In summary the variances or measures that were computed include largest mean variance (comparison of treatment means), largest contrasts variance, average mean variance and average contrasts variance.

Simulation studies

In the simulations three replications of treatments were assumed. The simulation procedures given in Section 4.3 are used except that for each simulation of $\mathbf{y} = \mathbf{f}^{\mathrm{i}} + \mathbf{t} + \mathbf{e}^{*}$, $\mathbf{y}^{r} = \mathbf{f}^{\mathrm{i}} + \mathbf{x}$ was also simulated alongside it. Notice that $\mathbf{y}^{r}$ is a vector without trend or correlation in the data points. The data vector $\mathbf{y}^{r}$ represented data from an RCBD. The computer simulations were used basically because they offer a valuable and feasible alternative to empirical tests of design efficiency. The main objective of these studies is to illustrate the statistical efficiency of systematic designs relative to RCBD. All designs I, II and III were evaluated. The relative efficiency of design I for two dimensional spatial variability was only evaluated under AR(1)×AR(1) covariance structure. For each and every setting 400 samples were generated.

The effects of the strength and the nature of the global trend on design II efficiency were also evaluated for spherical ($a = 6$) and AR(1) (at $\rho = 0.6$) spatial models. The values $a = 6$ and $\rho = 0.6$ were chosen because they had approximately average relative efficiency for spherical and AR(1) covariance structures, respectively. The linear trend was simulated using $\mathbf{t} = K(0.9 + 0.7R_{(i)})$ and that of quadratic trend using $\mathbf{t} = K(0.9+0.7R_{(i)} - 0.3R^{2}_{(i)})$ where $R_{(i)}$ is as defined in the previous chapter and K is a factor chosen to define trend strength. The settings used were 0.5, 1.0 and 1.5. Design II was chosen for illustration purposes.

The simulated data are analysed using among other statistical software packages PROC MIXED in SAS. The SAS PROC MIXED allows specification of the contrast matrix using CONTRAST statement. The effect of different values of the range param-

eter on efficiency was investigated.

Results are presented in Tables 5.1, 5.2, 5.3 and Table 5.4. In Table 5.1 and 5.3, efficiency is based on variances of mean differences evaluated under spherical and AR(1), and Gaussian and exponential models, respectively. In Table 5.2 and 5.4, efficiency is based on the variances of contrasts evaluated under spherical and AR(1), and Gaussian and exponential models, respectively. In both cases presented, the estimation is based on solely covariance structure modelling. The figures in bold in these tables are means for the sub-columns they immediately appear under. General results are presented in Table 5.5 where the trend effect is reported. Also included in Table 5.5 are other measures of efficiency and power to be described in the subsequent sections. The relative efficiencies for contrasts in Table 5.5 are the means from Tables Tables 5.1, 5.2, 5.3 and 5.4. Generally, results for spherical, Gaussian and exponential error processes had a common pattern. The assumption of one dimensional spatial variability for designs I, II and III gave the same results.

Discussion of simulation results

From the results in Tables 5.1, 5.2, 5.3 and Table 5.4, and 5.5 the following conclusions can be drawn;

1. Clearly, the relative efficiency decreases with the increase of the range parameter. This is due to many observations being correlated. It can also be concluded for the presence of an AR(1) error covariance structure that the stronger the correlation, the lower the efficiency of the design. Notice that large a implies many treatment levels and thus relatively large block size that may play a role in increasing the associated variance.

2. The relative efficiencies based on contrast variances are generally slightly lower than those based on mean variances. This is because the variance of contrasts is

Table 5.1: Relative efficiency of systematic designs to RCBD

Error model	Spatial parameter	Variance of means		Rel. effic. to RCBD	
		Large	Average	*for Large*	*for Average*
Spherical	a = 2	0.5604	0.4112	69.63	71.32
	a = 4	0.4388	0.3760	69.58	70.43
	a = 6	0.6276	0.6150	66.86	66.84
	a = 8	0.5735	0.5635	66.49	66.80
				68.14	**68.85**
AR(1)	$\rho = 0.3$	0.5693	0.5569	88.98	88.36
	$\rho = 0.6$	0.5822	0.5763	73.73	74.56
				81.36	**81.46**
AR(1)×AR(1)	$\rho = 0.3$	0.6194	0.6016	80.64	79.80
	$\rho = 0.6$	0.6239	0.6146	67.28	69.91
				75.46	**75.86**

more sensitive to spatial variation than variance of means.

3. Generally, there is an increase in the average variances of means and contrasts when the trend is present Table 5.5 (page 104). This is largely due to variance of treatment means being inflated by global trend variance. Thus, the efficiency of the design is reduced. This observation is in line with the argument that presence of a trend reduces efficiency of treatment comparisons (see Chapter 3). From Table 5.6 (page 105), increasing the trend parameter, K, tends to increase the average variance of contrasts. This implies a reduction in relative efficiency. However, this increase in variance is not significant when the trend is correctly specified or smoothed out. This supports the trend modelling approaches. Thus, from

Table 5.2: Relative efficiency of systematic designs to RCBD

Error model	Spatial parameter	Variance of contrasts		Rel. effic. to RCBD	
		Large	Average	*for Large*	*for Average*
Spherical	a = 2	0.5892	0.5803	68.36	68.86
	a = 4	0.6659	0.6600	61.80	63.33
	a = 6	0.6202	0.6091	67.20	67.43
	a = 8	0.6288	0.6002	64.73	64.98
				65.53	**66.15**
AR(1)	$\rho = 0.3$	0.4326	0.3961	83.436	84.01
	$\rho = 0.6$	0.4682	0.4466	76.46	75.86
				79.95	**80.00**
AR(1)×AR(1)	$\rho = 0.3$	0.6332	0.6234	79.64	79.80
	$\rho = 0.6$	0.6144	0.6244	72.05	69.91
				75.85	**74.86**

Table 5.6 (page 96), it can be concluded that trend strength may have no effect on treatment contrasts if the trend is correctly fitted. It was observed however, that if the trend is under-fitted the variances of contrasts are significantly larger than in well specified trend model. This is mainly for data with trend alone, otherwise error modelling would partly 'soak up' the remaining trend. There are no significant differences in the effect of nature of the trend.

4. The relative efficiencies are all well above 60% and on average 72%. Thus, in all cases, RCBD is superior to systematic designs even if error structures and trends are modelled. In other words, there is a statistical efficiency penalty due to use of the systematic designs. Note that the 60% efficiency is for comparing the main

Table 5.3: Relative efficiency of systematic designs to RCBD

Error model	Spatial parameter	Variance of means		Rel. effic. to RCBD	
		Large	Average	*for Large*	*for Average*
Exponential	a = 2	0.4930	0.4700	63.24	65.60
	a = 4	0.5685	0.5150	61.73	60.04
	a = 6	0.5380	0.5080	62.44	67.89
				62.47	**64.51**
Gaussian	$a = 2$	0.5085	0.4887	81.39	82.86
	$a = 4$	0.5931	0.5839	77.40	82.40
	$a = 6$	0.5869	0.5389	75.82	75.37
				78.87	**80.21**

Table 5.4: Relative efficiency of systematic designs to RCBD

Error model	Spatial parameter	Variance of contrasts		Rel. effic. to RCBD	
		Large	Average	*for Large*	*for Average*
Exponential	a = 2	0.6832	0.6234	64.36	64.64
	a = 4	0.6436	0.6241	62.31	60.30
	a = 6	0.6144	0.6143	64.24	64.20
				63.63	**63.05**
Gaussian	$a = 2$	0.4460	0.3942	81.31	80.36
	$a = 4$	0.4487	0.4230	82.64	84.23
	$a = 6$	0.5947	0.5697	69.57	72.09
				77.84	**78.89**

Table 5.5: Summary of relative efficiencies and relative power of systematic designs

Error model	Trend present	Efficiency for		Other measures		
		Means	Contrasts	Power	D_1effic.	D^seffic.
AR(1)		81.46	80.00	0.954	93.77	98.23
AR(1)	√	79.43	79.52	0.943	93.73	97.05
AR(1)×AR(1)		75.86	74.86	0.944	93.12	96.84
Exponential		64.51	63.05	0.959	88.63	95.67
Exponential	√	62.49	60.78	0.896	86.25	93.04
Gaussian		80.21	78.89	0.937	95.36	99.32
Gaussian	√	68.32	67.53	0.937	93.08	97.43
Spherical		68.85	66.15	0.971	89.91	98.59
Spherical	√	63.34	63.48	0.973	87.52	97.07
iid	√	92.62	92.63	0.953	98.03	98.04

plot treatments. However, there is a counter argument in practice to this penalty. In practice i.e. field conditions, the field trend or spatial variability will affect both RCBD and systematic designs in the same way although the effect may be more severe for systematic designs. This is based on the fact that in both classes of designs blocking is applied. In such instances, the efficiency of RCBD will be less than 100% and this will imply an increase in the relative efficiency of the systematic designs. For example consider the presence of a field trend and assume it affects the RCBD in the same way as it does a systematic design. From Table 5.5 the efficiency of the RCBD will be reduced to 92.62% for means or 96.63% for contrasts. Now consider the lowest efficiency in Table 5.5 i.e. 60.78% associated with exponential error correlation. In line with the above arguments this value

Table 5.6: Effect of nature and strength of global trend on variance of contrasts and D^s efficiencies under spherical (a = 6) and AR(1) ($\rho = 0.6$) covariance structures

Error model	Strength of trend (K)	Nature of the global trend	Average contrast variance	Relative efficiency	Power
Spherical	0.5	linear	0.5716	66.34	0.903
	0.5	quadratic	0.5798	65.91	0.863
	1.0	linear	0.6022	66.16	0.903
	1.0	quadratic	0.6051	66.31	0.915
	1.5	linear	0.6047	63.02	0.914
	1.5	quadratic	0.6038	63.13	0.903
AR(1)	0.5	linear	0.4479	70.13	0.929
	0.5	quadratic	0.4490	70.29	0.912
	1.0	linear	0.5321	68.94	0.893
	1.0	quadratic	0.5316	68.48	0.905
	1.5	linear	0.5214	69.14	0.903
	1.5	quadratic	0.5321	69.10	0.912

can be updated to 65.62%, a 3.6% increase.

5. The relative design efficiency depends on the error/covariance structure present in the field. The efficiency is highest if AR(1) (80% on average for contrasts) is present and lowest when exponential error correlations are present (62% on average for contrasts).

6. Design I has lower efficiency compared to design II and III if there is existence of two dimensional spatial variability in the observations exists. This is evidenced by the differences in efficiency results for AR(1) for designs II and III in comparison

with AR(1)×AR(1) for design I (Table 5.5, page 104).This is principally due to reduction in design efficiency further along the column from that given by design II and III. The efficiency given by design II and III can be viewed as design efficiency affected by spatial variability along the rows. Otherwise, if one dimensional spatial variability is assumed, the three designs have the same relative efficiency.

5.3.2 Relative design power

The average statistical power with which a particular hypothesis H is tested under the systematic design relative to randomised designs was computed. Power is defined as the probability of detecting departures from a null hypothesis of interest i.e. the probability that the decision rule will lead to conclusion of H_a when in fact H_a holds. The main idea is that from Pitman's efficiency equation (5.6), it is clear that hypothesis testing is directly related to design efficiency. Hypothesis testing, for example, can be done using the generalized F-statistic given in equation (4.24) which is distributed approximately as $\mathbf{F}_{[\text{rank}(L),v,\lambda_L]}$. The non-centrality parameter, λ_L, is given by

$$\lambda = (L'\boldsymbol{\beta})'\left[L'(X'H(\boldsymbol{\theta})^{-1}X)^{-1}L\right]^{-1}(L'\boldsymbol{\beta}) \tag{5.8}$$

Under $H_0 : \lambda_L = 0$ ($\mathrm{L}'\boldsymbol{\beta} = 0$) but when H_0 is false then $\lambda_L > 0$. The exact value of λ_L depends on $L'\boldsymbol{\beta}$, X and $H(\boldsymbol{\theta})$, where these terms are the magnitude of departure from H_0, the design and associated replication, and the variance and covariance components, respectively. Power is given by expression

$$P\left(\mathbf{F}_{[\text{rank}(L),v,\lambda_L|} > \mathbf{F}_{crit}\right) \tag{5.9}$$

where λ_L is the value of the non-centrality parameter under the alternative hypothesis of interest and it is a measure of how unequal the treatment contrasts or means are while $\mathbf{F}_{crit} = \mathbf{F}_{[\text{rank}(L),v,0,\alpha]}$ is the value of central $\mathbf{F}$ at the given α-level. A design that maximises the power of the test is desired. Thus, given equal replications of treatments

in designs ξ_S and $\xi_{\mathbb{R}}$, and testing the same hypothesis $L'\beta = 0$, their power relative to each other can be computed. Since variance and covariance components are defined by a particular design in question, this reduces to comparing the designs. In this sense power is directly related to design efficiency.

Simulation studies

The data simulated in Section 5.3.1 was used to compute design power. The power of the designs was computed using CONTRAST statement of PROC MIXED and a necessary program written in SAS. Results are presented in Table 5.5 (page 104).

1. The power of systematic designs is on average 94.2% (0.942) that of RCBD. This shows that although the ϕ_p directed efficiencies are low, the power of hypothesis testing in systematic designs is almost as good as that of RCBD. Contrary to ϕ_p directed efficiencies, the power does not vary much according to the error structure or trend present in the data. Possibly because the modelling of the trend or error structure precludes their strong influence on comparison of treatments. This is an assurance that the same conclusions for systematic design and RCBD can be arrived at within 10% confidence interval.

2. From the above results, systematic designs perform well under hypothesis testing. This at first sight appears to be contradictory i.e. measures of efficiency conflict. However, it can immediately be recognised that different aspects of an experiment are involved in these measures. A design can be optimum for one investigation but not for the other. For instance, it may be optimal for hypothesis testing and not optimal for point estimation of parameters (Kiefer, 1959). The systematic designs are as almost efficient as RCBD under hypothesis testing about main plot treatments, their interactions and also their interactions with split plot treatments. Notice that systematic designs I, II and III have almost the same relative pow-

ers of hypothesis testing. This again is indicated by the results associated with AR(1) and AR(1)×AR(1) covariance structures. The reason could be that error modelling precludes ifs influence on the power of the F test.

5.3.3 D-efficiency for regression parameters

Consider fitting a polynomial regression model (beta-hat model II) with a prime interest in defining the shape or pattern of the response. For efficient evaluation of this response surface the regression coefficients or parameters should be estimated as precisely as possible. The efficiency of different designs can be evaluated based on how precise these coefficients are estimated. Efficiency measures such as D_1 and D^s can be used to compare systematic designs.

D^s efficiency

The standardised D-efficiency, D^s, of an arbitrary design ξ with respect to a D-optimal design ξ^* is defined as

$$I(\xi) = \left\{ \frac{|M(\xi)|}{|M(\xi^*)|} \right\}^{\frac{1}{m}} \tag{5.10}$$

where m is the number of the parameters in the model, say $m = 2$, 3 or 4. In intercropping plant population studies the second order (m = 2) and third order (m = 3) parameters are very important. Thus for estimation of the optimum or proper evaluation of the response surface/pattern efficient estimation of these parameters is very important. Since the denominator is constant in (5.10) the relative efficiency for each parameter for systematic design ($\xi_{\mathbb{S}}$) with respect to RCBD ($\xi_{\mathbb{R}}$) is given by

$$I(\xi_{\mathbb{S}}) = \left\{ \frac{|M(\xi_{\mathbb{S}})|}{|M(\xi_{\mathbb{R}})|} \right\}^{\frac{1}{m}} \tag{5.11}$$

Observe that the asymptotic variances for the regression parameters are given as inverse of their information matrices. Thus the D^s efficiency for each parameter can be defined

in terms of its asymptotic variance, i.e.

$$D_i^s = \frac{\left(M^{-1}(\xi_{\mathbb{S}})\right)_{ii}}{\left(M^{-1}(\xi_{\mathbb{R}})\right)_{ii}}, \qquad i = 1, 2, 3. \tag{5.12}$$

The equation (5.12) provides the basis for evaluating the efficiency of systematic designs to RCBD in this study.

D_1 efficiency

The D_1 efficiency of an arbitrary design ξ with respect to a D-optimal design ξ^* is defined as

$$eff_m^{D_1}(\xi) = \left(\frac{|M_m(\xi^*)|}{|M_{m-1}(\xi^*)|}\right) \Big/ \left(\sup_{\xi} \frac{|M_m(\xi)|}{|M_{m-1}(\xi)|}\right) \tag{5.13}$$

The D_1 efficiency maximises the power of the t-test or F-test for the significance of the highest coefficient in the polynomial of degree m i.e. for the hypothesis that $H_0 : \beta_m = 0$ (Dette and Franke, 2000) where β_m is a regression parameter of polynomial degree m. It thus minimises the variance of the estimator for the coefficient β_m. Again in the present study the denominator in (5.13) is a constant and thus the relative D_1 efficiency can be computed by considering the numerator i.e.

$$\frac{|M_m(\xi^*)|}{|M_{m-1}(\xi^*)|} \tag{5.14}$$

The critical region for testing the hypothesis $H_0 : \beta_m = 0$ is given by

$$F \geq F_{crit}$$

and

$$P(F \geq F_{crit})$$

gives the power function of this critical region. Noting that the expression (5.14) has a central F distribution under H_0 and a non-central F distribution under H_A, the non-centrality parameter plays a major role in definig the power of F-test. This leads to

comparative relative design power like that in Section 5.3.2 except in this case power is specifically for regression parameters.

Simulation studies

The same simulations as in the previous sections were used. The D_1 and D^s efficiencies were defined for linear and quadratic functions of crop densities. Regression analysis was implemented using PROC MIXED in SAS (some implementations were undertaken using REML in Genstat and slmr in SPLUS). Results for $m = 2$ are presented in Table 5.5 (page 97). Some additional results for $m = 1$ (linear effect) and $m = 2$ (quadratic effect) are given in Table 5.7.

Table 5.7: Summary of standard errors and power for β_m; $m = 1, 2$

Error	s.e for $m = 1$		s.e for $m = 2$		Power for $m = 2$	
model	$\xi_{\mathbb{S}}$	$\xi_{\mathbb{R}}$	$\xi_{\mathbb{S}}$	$\xi_{\mathbb{R}}$	$\xi_{\mathbb{S}}$	$\xi_{\mathbb{S}}$
AR(1)	0.2531	0.2305	0.0334	0.0328	0.5137	0.5229
Exponential	0.3888	0.3602	0.0422	0.03926	0.6301	0.6587
Gaussian	0.3167	0.3165	0.0393	0.0389	0.8010	0.8065
Spherical	0.2437	0.2221	0.0374	0.0367	0.6835	0.6972

1. The standard errors associated with regression parameters/ coefficients computed under assumptions of designs I, II and III were considerably the same.

2. The relative D_1 and D^s efficiencies (presented in Table 5.5, page 97)) of systematic designs are on average above 89%. This indicates the merit of systematic design and the opportunity of using them since regression parameters are well estimated. The lowest relative D_1 efficiency is 86.25% and the lowest D^s efficiency is 93.04% all associated with spatial variability involving exponential covariance structure

and the global trend. Notice that if the same argument as in Section 5.3.1 is employed (that is spatial variability affects systematic designs and RCBD in a similar way), some relative efficiencies of systematic designs will be higher than 100% indicating superiority of systematic designs over RCBD. For example, consider D^s for Gaussian covariance structure (99.32%) and a corresponding 98.4% for RCBD. The updated relative efficiency of systematic designs will be 101.3%.

3. Observe that the s.e's associated with regression parameters (both $m = 1$ and $m = 2$) for systematic design and RCBD are close to one another in all the covariance models. Thus, modelling error structure in systematic designs yields precise estimates of regression parameters. It was observed that in more than 23% of simulated data samples, systematic design gave higher precision of regression parameters than RCBD.

5.4 Efficiency under cost function

Given the advantages of involving many plant populations in preliminary plant population studies, the cost of conducting such factorial experiments is of great importance. In addition to costs incurred in non-factorial experiments, factorial experiments incur extra costs due to changing the levels of each of the factors. In this case, changing a plant population level will involve a cost associated with introduction of a guard row. Thus, if cost C_i is attached to a guard row the total cost can be defined in terms of level changes.

Definition 5.4 For any factor A_j, $j = 1, 2, \ldots, n$, in the design ξ, the cost for the factor is defined as;

$$C(A_j) = \sum_{i=1}^{r} C_i(A_j)$$

Where $C_i(A_j)$ is the number of times of level changes of factor A_j in the ith block/main plot i.e. is the number of level changes in $\mathbf{X}_j$, design matrix of changes in A_j, except for the changes between main plots and blocks. As an example, consider a level sequence of a factor in a three-level factorial design given as follows;

$$(i)\begin{pmatrix}0\\1\\2\end{pmatrix} \qquad (ii)\begin{pmatrix}0\\2\\1\end{pmatrix} \tag{5.15}$$

Although the level changes of the level of sequence (5.15(i)) is equal to 2, there will be no cost incurred. This by definition of systematic designs being considered in this study requires no guard rows. Consider factorial design with sequence in (5.15(ii)). The level changes in this sequence is again equal to 2. This is a randomized sequence and by definition of randomized designs, two guard rows are required. The cost is therefore incurred due to introduction of guard rows. It will be shown in the subsequent paragraphs that this cost limits on the number of practical replicates in a randomized design in situations where land area is limited.

Another concept of comparison of systematic designs to randomised designs is the harvestable land area. Normally, in field experiments it is a portion of the plot which is harvested and not the whole plot, the argument being that there may be some carryover of treatment effects from plots to neighbouring plots. That is a proportion $s_i\,(i = 1, 2)$ is harvested and this proportion can be thought of as a sample size where s_1 is the proportion of plot harvested in randomised design while s_2 is the proportion harvested in a systematic design. In systematic designs a larger size of a plot area is harvested compared to randomised designs where only the inner rows in a plot are harvested. Mead and Riley (1981) and Mead (1994) working empirically have shown that s_2 is about 80% while s_1 is below 60%. Also in the 3 experiments described in Chapter 2, the proportions of total harvestable area were 59.3%, 61.0% and 63.8%.

For r replicates the information on each treatment mean assuming the variance components are known is

$$\frac{rs_i}{\sigma^2_{s_i} + s_i\sigma^2_e} \tag{5.16}$$

where σ^2_e is error variance and $\sigma^2_{s_i}$ is the variance associated with the proportion s_i with the definition constraint

$$\lim_{s_i \to 1} \sigma^2_{s_i} = 0$$

Since $s_2 > s_1$ the information from s_2 is expected to be greater than that from s_1.

The land area required by a split plot in RCBD and harvestable land area from the same relative to a parallel-row systematic design were computed. The computations were based on secondary data from technical reports at research institutes in Uganda and theses at Makerere university concerning ten intercrop studies. Although these studies were carried out using split plot in RCBD, the area that would be required by the parallel-row systematic design was approximated by assuming absence of guard rows. In other words the figures presented in Table 5.8 were calculated as a ratio of total land area taken by intercrop in split plot design and the would-be area if guard rows were excluded (to represent a systematic design). The relative harvestable area figures were calculated basing on the assumption the parallel-row systematic design would allow more than 60% of each plot to be harvested. The intercrops considered include cowpea-maize, maize-beans, simsim-maize, sorghum-simsim, sorghum-cowpea, cassava-maize, Gnuts-maize, simsim-finger millet and finger millet-pigeon peas. The figures presented are averages over the area occupied by these intercrops. From the table, it is clear that a parallel-row systematic design is more efficient in utilising land compared to a split plot design.

Thus, on average for the same experiment RCBD will require 40% more land compared to systematic designs. Also on average 80% more land area is harvested in the case of systematic design compared to RCBD.

Table 5.8: Ratio of land area of randomized block to systematic block

Experiment size	Required land area	Harvestable land area
6×6	1.33	1.63
8×6	1.25	1.77
8×8	1.40	1.94
6×12	1.63	1.83

Consider the design ξ to have a cost $K(\xi)$ continuous on space χ and impose a general constraint that the expected total cost of experimentation should not exceed K_0. Thus,

$$\int_{\chi} K(\xi) d\xi \leq K_0$$

The problem is then to maximize $M(\xi*)$ in some sense subject to this constraint.

$$M(\xi*) = \frac{M(\xi)}{K(\xi)}$$

Cost can be viewed as defining the number of replication or plot sizes to be used in an experiment i.e. a function of the number of replications and or plot sizes. Assuming a unit cost for each m^2 of land, fixing the cost of experimentation at K_0 is equivalent to fixing amount of land area available. This would lead, for the same land area, to fewer replications in RCBD compared to systematic designs. The $K(\xi)$ is then a diagonal matrix with its elements being a ratio of replicates in systematic design to a randomized design.

Example 5.4.1 Consider an experiment consisting of 2 spatial arrangements for 3 plant densities of maize and 6 beans plant densities laid out in a split plot or RCBD with plot sizes 5×4 m. Each replicate would consist of 36 plots and there will be 30 level changes. In this case, approximate land area under every 4 guard rows is equivalent to

area under a full plot. In total, an area equivalent to 10 plots would be available if a parallel row design is used. Overall, therefore, for a land area that allows 3 replicates in an RCBD, a parallel row design would allow 4 replicates.

Example 5.4.2 Consider an experiment with 2 spatial arrangements for 4 plant densities of sorghum and 6 cowpea plant densities laid out in a split plot with plot sizes 6×5. Each replicate consists of 48 plots and if a given land area is enough for 5 replicates using a parallel row design, only 4 replicates are possible with an RCBD.

Example 5.4.3 Consider an experiment with 2 spatial arrangements for 4 maize plant densities and 8 cowpea plant densities laid out in a split plot with plot sizes 6×5 m. Each replicate consists of 32 plots and if a given land area is enough for 4 replicates using a parallel row design, only 3 replicates are allowable with an RCBD.

From equation (5.16), the information on each treatment mean would be lower under fewer replications. Thus using a cost function constraint is like evaluating the effect of replication on the efficiency of the design. The ϕ_p directed efficiencies and relative power of systematic design were revaluated under this cost function.

Presented in Table 5.9 are the results based on four replications for 200 realisations for each setting of trend and covariance models. There is an improvement in the statistical relative efficiencies for systematic designs. The lowest relative efficiency in the table being 64.42% compared to 60.78% associated with the exponential covariance model previously under three replicates. Whereas reasoning out cost in terms of replicates may be misleading, it is still informative about the relative efficiency of systematic designs.

In closing this chapter, it is noted that systematic designs generally are efficient designs compared to RCBD. Since the main purpose for population studies in intercropping experiments is centered on response surface pattern, their use is very reliable

Table 5.9: Relative efficiencies and Power of systematic designs for four replications

Error model	Trend present	Efficiency for Means	Efficiency for Contrasts	Relative Power
AR(1)		83.55	82.10	0.961
AR(1)	√	82.63	80.27	0.954
AR(1)×AR(1)		80.58	79.41	0.924
Exponential		64.82	64.42	0.987
Exponential	√	64.06	63.43	0.949
Gaussian		84.89	84.91	0.960
Gaussian	√	73.41	73.04	0.945
Spherical		70.98	70.48	0.967
Spherical	√	68.32	69.32	0.897
iid	√	94.85	94.86	0.986

and efficient as compared to RCBD. The additional advantages make systematic designs more preferable to RCBD in plant population studies.

Chapter 6

Conclusions

This study has been on exploration and evaluation of spacing systematic designs with respect to intercropping studies involving plant populations. The exploration and development of methods of analysis and modelling of data and evaluation of efficiency of these spacing systematic designs in intercropping experiments formed the basis of this work. In this study the advantages of systematic designs have been elaborated through a literature review and theoretical developments. It is shown that systematic designs have a major role to play in intercropping population studies and examples of scenarios where they are effectively applicable are given. The conceptual framework of the study has been based on the existence of spatial variability in the data from systematic designs. Existence of spatial variability has been the main argument against use of systematic designs. In this study, it is shown that spatial variability is easily modelled.

It has been shown both through theoretical arguments in Chapter 3 and simulations in Chapter 4 that the effect of spatial variability can easily be handled. The use of trend and error modelling techniques such as spatial linear mixed models, semi-parametric mixed models and beta-hat models incorporating spatial variability have shown to be effective. Therefore apart from offering advantages such as allowing many treatment

levels to enable effective exploration of response pattern and requiring small land area compared to RCBD, systematic designs can be efficiently analysed.

The beta-hat models introduced have been shown to perform well in handling spatial data from systematic designs. Since these models combine the information over systematic main plots and enable comparing them, they provide a powerful tool for handling data from intercropping plant population studies. Through simulations or numerical calculations these methods have been shown to be reliable and efficient.

Certain conclusions concerning the validity of the methods seem justified. Compared to correctly fitting the global trend, over-fitting trend had little effect on the validity of error modelling in the case where both trend and error correlation were present. Under-fitting trend affects error modelling but had little effect on the validity of tests and estimates of precision. The use of smoothing splines particularly provides a flexible framework for identifying and modelling the trend. These methods are powerful and are easily implemented. The use of the real field data set illustrates that the general spatial mixed models and beta-hat models perform in a similar and satisfactory way.

There is more to be gained by good estimation of response pattern parameters. Through simulations in this study it has been shown that systematic designs are almost as efficient (93%) as RCBD in estimation of these parameters. With additional practical advantages as argued out in Chapter 3 and Chapter 5, it is reasonable to conclude that systematic designs provide the best alternative for preliminary plant population studies. Besides this, systematic designs have been shown in general to enable comparing main plot treatments with 72% relative efficiency. This is reassuring and guarantees correct conclusions. All the three systematic designs scenarios (i.e. design I, II and III) are equally good for preliminary plant population investigations.

It has been shown that modelling the intercropping effect as an additive effect to sole

crop yield leads to same the conclusions as modelling yield itself. It has also been shown theoretically that modelling of relative LER and ER in terms of plant populations is also equivalent and similar in handling and modelling to crop yields.

The contributions of this study therefore has been two fold:

- the application of appropriate statistical analysis and modelling systematic design data. It has been shown that trend modelling and error modelling are powerful, reliable and efficient tools. They are also easily implemented. Thus, use of designs that precludes error or trend modelling is not a strong prerequisite in efficient experimentation anymore.

- the definition of systematic design efficiency. In the interest of response pattern exploration and given cost advantage over RCBD, systematic designs provides an efficient and reliable alternative. This is especially in early stages of intercropping plant studies.

Further theoretical developments are needed to gain insight into the following

- The use and validity of spatial modelling techniques in bivariate analysis of intercropping responses. Also research into other multivariate data handling methods incorporating spatial variability is needed.

- The close relationship between ARMA(1, 1) and spherical covariance models needs further investigation

- The distribution of intercropping additive effects and ER needs further research.

- The incorporation of covariates into the suggested models is yet an area requiring investigation.

Bibliography

Azzalini, A. and Giovagnoli, A. 1987. Some optimal designs for repeated measurements with autoregressive errors. *Biometrika* **74**: 725 – 734.

Bailey, R.A., Azaïs, J.M. and Monod, H. 1995. Are neighbour methods preferable to analysis of variance for completely systematic designs? "Silly designs are silly?" *Biometrika* **82**: 655 – 659.

Baird, D. and Mead, R. 1991. The empirical efficiency and validity of two neighbour models. *Biometrics* **47**: 1473 – 1487.

Balaam, L.N. 1986. Intercropping:– past and present. In *Statistical Design: Theory and Practice. Proceedings of a Conference in Honour of Walter T. Federer* (C.E. McCulloch, S.J. Schwager, G. Casella, and S.R. Searle, eds.). Cornell University, Ithaca, New York. pp141 – 148.

Bartlet, M.S. 1978. Nearest neighbour models in the analysis of field experiments. *Journal of Royal Statistical Society B* **40**: 147 – 174.

Besag, J. and Kempton, R. 1986. Statistical analysis of field experiments using neighbouring plots. *Biometrics* **42**: 231 – 251.

Bleasadale, J.K.A. 1967. Systematic designs for spacing experiments. *Experimental Agriculture* **3**: 73 – 85.

Brownie, C. and Gumpertz, M.L. 1997. Validity of spatial analysis for large field trials. *Journal of Agricultural, Biological and Environmental Statistics* **2**: 1 – 23.

Brownie, C., Bowman, D.T. and Burton, J.W. 1993. Estimating spatial variation in analysis of data from yield trials: A comparison of methods. *Agronomy Journal* **85**: 1244 – 1253.

Cochran, W. G. and Cox, G.M. 1964. *Experimental Design.* Second edition. Wiley, New York.

Cox, D.R. 1951. Some systematic experimental designs. *Biometrika* **38**: 313 – 323.

Cressie, N. 1991. *Statistics for Spatial Data.* John Wiley, New York.

Cressie, N.A.C., and Hartfield, M.N. 1996. Conditionally specified Gaussian models for spatial statistical analysis of field trials. *Journal of Agricultural, Biological and Environmental Statistics* **1**: 60 - 77.

Cullis, B.R. and Gleeson, A.C. 1991. Spatial analysis of field experiments – an extension to two dimensions. *Biometrics* **47**: 1449 – 1460.

De Wit, C.T. and Van den Bergh, J.P. 1965. Competition between herbage plants. *Netherlands Journal of Agricultural Science* **13**: 212 – 221.

Dette, H. and Franke, T. 2000. Constrained D- and D_1- optimal designs for polynomial regression.

Durban, M., Hackett, C. and Currie, I. 1997. Approximate standard errors in semiparametric models. *Biometrics***55**: 699 – 703.

Fairfield Smith, H. 1938. An empirical law describing heterogeneity in the yields of agricultural crops. *Journal of Agricultural Science* **20**: 339 – 351.

Federer, W.T and Schwager, S.J. 1982. On the distribution of land equivalent ratios. Unpublished Technical Report BU – 777 – M in the Technical Report Serie of the Biometrics Unit, Cornell University, Ithaca, New York.

Federer, W.T. 1993. *Statistical Design and Analysis for Intercropping Experiments.* Vol. I: Two Crops. Springer Series in Statistics, New York.

Federov, V. and Müller, W. 1989. Comparison of two approaches in optimal design of an observation network. *Statistics* **20**: 339 – 351.

Finney, D.J. 1988. Was this your statistical text book? III. Design and Analysis. *Experimental Agriculture* **24**: 421 – 432.

Freshman, A., Cullis, B.R. and Verbyla, A.P. 1997. Genotype by environment variance heterogeneity in a two stage analysis. *Biometrics* **53**: 1373 – 1383.

Gilmour, A.R., Cullis, B.R. and Verbyla, A.P. 1997. Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological and Environmental Statistics* **2**: 269 – 293.

Gleeson, A.C. and Cullis, B.R. 1987. Residual maximum likelihood (REML) estimation of a neighbour model for field experiments. *Biometrics* **47**: 277 – 288.

Green, P.J. and Silverman, B.W. 1994. *Nonparametric Regression and Generalised Linear Models.* Chapman and Hall, London.

Grondona, O.M, and Cressie, N. 1991. Using spatial considerations in analysis of experiments. *Technometrics* **33**: 381 - 391.

Harris, R.T and Dallas. E.J. 1996. A regression model with spatially correlated errors for comparing remote sensing and In-Situ measurements of a grassland site.

Journal of Agricultural, Biological and Environmental Statistics **2**: – 23.

Hastie, T.J. and Tibshirani, R.J. 1990. *Generalised Additive Models.* Chapman and Hall, London.

Henderson, C.R. 1984. *Applications of Linear Models in Animal breeding.* University of Guelph.

Henderson, C.R. 1990. Statistical methods in animal production improvement: Historical Review in: *Advances in Statistical Methods for Genetic Improvement of Livestock.* Springer - Verlag. pp 1 - 14.

Huxley, P.A. and Maingu, Z. 1978. Use of systematic spacing design as an aid to the study of intercropping: Some general considerations. *Experimental Agriculture* **14**: 49 – 56.

Jensen, D.R. and Ramirez, D.E. 1993. Efficiency of comparisons in linear inference. *Journal of Statistical Planning and Inference* **37**: 51 – 68.

Journel, A.G and Huijbregts, C.J. 1978. *Mining Geostatistics.* London. Academic press.

Kempthorne, O. 1952. *The Design and Analysis of Experiments.* Wiley, New York.

Kenward, M.G and Roger, J.H. 1997. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics* **54**: 983 – 989.

Kiefer, J. C. 1959. Optimum experimental designs. *Journal of Royal Statistical Society, B* **21**: 273 - 319.

Kiefer, J. C. 1974. General equivalence theory for optimum designs. *Annals of Statistics* **2**: 849 – 879.

Kirk, H.J., Haynes, F.L. and Monroe, R.J. 1980. Application of trend analysis to horticultural field trials. *Journal of American Horticultural Science* **105**: 189 – 193.

Levin, J.B. 1999. Use of variance component models in the analysis of complex surveys. PhD Thesis, University of Natal. pp 143 - 146.

Lill, W.J., Gleeson, A.C. and Cullis, B.R. 1988. Relative accuracy of a neighbour method for field trials. *Journal of Agricultural Science Cambridge* **111**: 339 – 346.

Littel, R.C, Milken, G.A, Stroup, W. and Wolfinger, R.D. 1996. *SAS System for Mixed Models.*SAS Institute Inc. NC.

Martin, R.J. 1990. The use of time-series models in the analysis of agricultural field trials. *Communications in Statistics – Theory and Methods* **19**: 55 – 81.

Martin, R.J. 1996. Spatial experimental design. In: *Handbook of Statistics* **13**. Eds. Ghosh, S. and Rao, C.R. pp 477 – 514. Amsterdam, North Holland.

McCulloch, C.E. and Searle, S.R. 2001. *Generalised, Linear and Mixed Models.* John Wiley and Sons, Inc., New York.

Mead, R. 1970. Plant density and crop yield. *Applied Satistics 19*: 64 – 81.

Mead, R. 1979. Competition experiments. *Biometrics* **35**: 41 – 54.

Mead, R. 1980. Designing experiments for intercropping research. *Experimental Agriculture* **16**: 329 – 342.

Mead, R. 1994. *The Design of Experiments.* Statistical principles for practical application. Cambridge Univ. Press, Great Britain.

Mead, R. 1995. A review of analysis of experimental data. In: *Proceedings of the*

1995 Kansas State University Conference on Applied Statistics in Agriculture. Pp 1 – 20. KSU, Manhattan.

Mead, R. and Stern, R.D. 1980. Designing experiments for intercropping research. *Experimental Agriculture* **16**: 329 - 342.

Mead, R. and Riley, J. 1981. A review of statistical ideas relevant to intercropping research (with discussion). *Journal of Royal Statistical Society* A **144**: 462 – 509.

Mead, R. and Willey, R.W. 1980. The concept of land equivalent ratio and advantages in yields from intercropping. *Experimental Agriculture* **24**: 217 - 228p.

Milliken, G.A. and Johnson, D.E. 2002. *Analysis of Messy Data.* Vol. III: Analysis of Covariance. Chapman and Hall/CRC, New York.

Müller, W. 2000. *Collecting Spatial Data: Optimum Design of Experiments for Random Fields.* Second edition, Physica Verlag, Heidelberg.

Nelder, J.A. 1962. New kinds of systematic designs for spacing experiments. *Biometrics* **18**: 283 – 307.

Ocaya, P. CYP. 1998. Effect of time of planting and spatial arrangement on the growth and yield of cowpea intercropped with maize. MSc. Thesis, Makerere University, Kampala. 121pp.

Osiru, D.S.O. and Willey, R.W. 1972. Studies on mixtures of dwarf sorghum and beans (*Phaseolus vulgaris*) with particular reference to plant population. *Journal of Agricultural Science Cambridge* **79** (3): 531 – 540.

Papadakis, J.S. 1937. Methodé statistiqué pour des experiénces sur champs. *Bull. Inst. Amel. Plantes a Salonique,* 23.

Pearce, S.C. and Gulliver, B. 1979. The statistical analysis of data from intercropping experiments. *Agricultural Science Cambridge* **93**: 51 – 58.

Peterson, R.H. 1994. *Agricultural Field Experiments.* Design and Analysis Marcel Deker Inc. New York.

Patterson, H.D. and Thompson, R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika* **58**: 545-554.

Pulkelshem, F. 1993. *Optimal Design of Experiments.* Wiley, New York.

Robinson, G.K. 1991. That BLUP is a good thing: the estimation of random effects. *Statistical Science* **6**: 15 – 51.

SAS Insititute Inc. 2001. SAS/STAT software. *Changes and Enhancements through Release 8.2.* Cary, NC; SAS Institute Inc.

Steele, R.G. D. and Torrie, J.J.W. 1990. *Principles and Procedures of Statistics.* A Biometrical approach. 2^{nd} Edition.: McGraw Hill Book Company Inc. New York.

Verbyla, A.P. and Cullis, B.R. 1992. The analysis of multistratum and spatially correlated repeated measures data. *Biometrics* **42**: 1015 – 1032.

Verbyla, A.P., Cullis, B.R., Kenward, M.G. and Welham, S.J. 1999. The analysis of designed experiments and longitudinal data using smoothing splines (with discussion). *Journal of Royal Statistical Society C* **48**: 269 – 311.

Wahba, G. 2000. Smoothing Splines in Nonparametric Regression. *Technical Report no. 1024.* University of Wisconsin, Madison, WI 53706.

Wahua, T.A.T and Miller, D.A. 1978. Relative yield total and yield components of intercropped sorghum and soybeans. *Agronomy Journal* **70**: 287 – 291.

Wang, Y. 1998. Mixed effects smoothing splines ANOVA. *Journal of Royal Statistical Society B* **60**: 159 – 174.

Wang, Y. and Taylor, J.M.G. 1995. Inference for smooth curves in longitudinal data with application to AIDS clinical trial. *Statistics in Medicine* **14**: 1205 – 1218.

Watson, S. 2000. Spatial dependence and block designs in spaced plant herbage trials. *Journal of Agricultural Science Cambridge* **134**: 245 - 258.

Wijesinha, A. 1981. Contribution to a paper by Mead, R. and Riley, J. (1981): A review of statistical ideas relevant to intercropping research (with discussion). *Journal of Royal Statistical Society* A **144**: 462 – 509.

Wilkinson, G.N., Eckert, S.R., Hancock, T.W. and Mayo, O. 1983. Nearest neighbour (NN) analysis of field experiments (with discussion). *Journal of Royal Statistical Society* B **45**: 151 – 211.

Willey, R.W. 1979. Intercropping, its importance and research needs. Part II: Agronomy and research approaches. *Field Crops Abstracts* **32**: 73 – 85.

Willey, R.W. and Heath, S. 1969. The quantitative relationships between plant population and crop yield. *Advances in Agronomy* **21**: 281 – 321.

Willey, R.W. and Lakhani, D.A. 1976. Some aspects of the productivity and resource use of mixtures of sunflower and fodder radish. *In:* Symposium on intercropping in Semi-arid Areas Proceedings, Morogoro, Tanzania 1 – 10.

Willey, R.W. and Osiru, D.S.O. 1972. Studies on mixtures of maize and beans (*Phaseolus vulgaris*) with particular reference to plant population. *Journal of Agricultural Science Cambridge* **79** (3): 519 – 529.

Willey, R.W. and Rao, M.R. 1980. A competitive ratio for quantifying competition between intercrops. *Experimental Agriculture* **16**: 117 – 125.

Willey, R.W. and Rao, M.R. 1981. A systematic design to examine effects of plant population and spatial arrangement in intercropping. *Experimental Agriculture* **17**: 63 – 73.

Williams, E.R. 1986. Row and Column designs with contiguous replicates. *Australian Journal of Statistics* **28**: 154 – 163.

Williams, R.M. 1952. Experimental designs for serially correlated observations. *Biometrika* **39**: 151 – 167.

Willy, R.W. and Rao, M.R. 1981. A systematic design to examine effects of plant population and spatial arrangement in intercropping, illustrated by an experiment on chickpea/safflower. *Experimental Agriculture* **17**: 63 – 73.

Yates, F. 1939. The comparative advantage of systematic and randomised arrangements in the design of agricultural and biological experiments. *Biometrika* **30**: 440 – 466.

Zimmerman, D.L. and Harville, D.A. 1989. On the Unbiasedness of the Papadakis estimator and other non-linear estimators of treatment contrasts in field – plot experiments. *Biometrika* **76**: 253 – 259.

Zimmerman, D.L. and Harville, D.A. 1991. A Random Field approach to the analysis of field – plot experiments. *Biometrics* **47**: 223 – 239.

APPENDIX A

TABLE I

Analysis of variance

Variate: YSM (yield of simsim) (SpFM= density of finger millet, SpSM= density of simsim)

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Rep stratum	2	8305.	4153.	0.16	
Rep.Row stratum					
SpFM	4	597204.	149301.	5.58	0.019
Residual	8	213931.	26741.	8.48	
Rep.Row.Splt stratum					
SpSM	3	337533.	112511.	35.67	<.001
SpSM.SpFM	12	85211.	7101.	2.25	0.035
Residual	30	94634.	3154.		
Total	59	1336819.			

Variate: YSM

SpSM	SpFM	0.00	10000.00	20000.00	40000.00	60000.00
6330.00		352.2	215.6	260.0	162.2	213.3
11100.00		450.0	250.0	316.3	200.0	253.3
22200.00		488.9	260.0	387.4	238.5	286.0
44400.00		723.3	305.6	493.3	283.0	420.0

*** Least significant differences of means (5% level) ***

Table	SpSM	SpFM	SpSM SpFM
rep.	15	12	3
l.s.d.	41.88	153.95	166.40
d.f.	30	8	14.19
Except when comparing means with the same level(s) of			
SpFM			93.66
d.f.			30

TABLE II

Analysis of variance

Variate: RelLER

Source of variation	d.f.	s.s.	m.s.	v.r.	F pr.
Rep stratum	2	8343.	4171.	0.29	
Rep.Row stratum					
SpFM	3	88600.	29533.	2.05	0.209
Residual	6	86516.	14419.	1.33	
Rep.Row.Splt stratum					
SpSM	3	524612.	174871.	16.14	<.001
SpFM.SpSM	9	69105.	7678.	0.71	0.696
Residual	24	260034.	10835.		
Total	47	1037211.			

TABLE III

Variate: RelLER

SpFM	SpSM	6330.00	11100.00	22200.00	44400.00
10000.00		437.	582.	565.	593.
20000.00		412.	584.	617.	753.
40000.00		316.	423.	487.	661.
60000.00		408.	478.	572.	736.

*** Least significant differences of means (5% level) ***

Table	SpFM	SpSM	SpFM SpSM
rep.	12	12	3
l.s.d.	120.0	87.7	181.2
d.f.	6	24	27.99
Except when comparing means with the same level(s) of			
SpFM			175.4
d.f.			24

Appendix B

SAS simulation code for correlated data for testing **y** –E[**y**] term

```
data sample;
do samp = 1 to 400;

         rho = 0.3;

            /* first error term */
         eps = rho * rannor( 47392 ) + rannor( 82745 );

         do x = 1 to 150; /* to enable three replicates*/
            y = 2 + 5 * x + eps;
            eps=eps;
            eps = rho * eps  + rannor( 32815 );
            output;
         end;
      end;
```

GenStat code for simulation of data used in testing block effect example

```
for n = 1 to 600
      for mi = 121, 123,110,109,124,123, 100,110.5,129.3,
110.8,123.2,125.8 *chosen randomly*
       for R = 1,2...12
    GRANDOM [DISTRIBUTION=Normal; NVALUES=12; SEED=31245; MEAN=0;
VARIANCE=1] f
    calculate y = f+mi
              & Ti = 1.89 + 0.4 R + 0.18R**2

  calculate Y = y+Ti
        endfor
       endfor
 print Y
   endfor
```

Simulation codes for main simulations

```
Design I: AGHIERARCHICAL [Print=Design;Seed=32154]
BLOCKFACTORS=Block,Plots; TREATMENTFACTORS=*,!p(Treat1,Treat2,Treat3);
LEVELS=4,!p(8, 4,  3,10) *4 blocks/replicates, 10 levels of crop B*

Design II: AGHIERARCHICAL [Print=Design;Seed=41259]BLOCKFACTORS =Block,
WPlots, SPlots;
TREATMENTFACTORS=*,!p(W_Treat1,W_Treat2),S_Treat1;LEVELS=4,!p(4,2,)8


Design III: AGHIERARCHICAL [Print=Design;Seed=10070]
BLOCKFACTORS=Block, WPlots, S_Plots; TREATMENTFACTORS=*,!p
(W_Treat1,S_Treat); LEVELS=4,3,12)
```

(No randomization constraint was applied in design generation).

```
proc iml;
seed=124517;

V =
H = root(V); /* define the covariance matrix V E.g. AR(1) with ρ=0.3
for 10 plots is given as
V = {1 0.3  0.09  0.027 0.0081    0    0    0          0       0,
0.3   1     0.3   0.09  0.027 0.0081        0     0   0       0,
0.09  0.3   1     0.3   0.09  0.027 0.0081        0   0       0,
0.027 0.09  0.3   1     0.3   0.09  0.027 0.0081      0       0,
0.0081      0.027 0.09  0.3   1     0.3   0.09  0.027 0.0081 0,
0     0.0081      0.027 0.09  0.3   1     0.3   0.09  0.027 0.0081,
0     0     0.0081      0.027 0.09  0.3   1     0.3   0.09  0.027,
0     0     0     0.0081      0.027 0.09  0.3   1     0.3   0.09,
0     0     0     0     0.0081      0.027 0.09  0.3   1     0.3,
0     0     0     0     0     0.0081      0.027 0.09  0.3   1};

*/

W = T(H);

f1 = normal (repeat(257715,24))+ mean value;/* mean values chosen
randomly from a ranges such as 100 - 130, 50 - 58, etc.*/

n1 = normal (repeat(257715,24));
n2 = normal (repeat(232715,24));
n3 = normal (repeat(259915,24));
n4 = normal (repeat(957715,24));

y1 =  f1 + W*n1;
do i = 1 to 3;
print y1;
end;
quit;
```

Example generating data for design I

Factor levels: a = b = 10; p = 1(fixed spatial arrangement). Parameters: $\rho = 0.3$. Design of 10 rows and 10 columns per block with 3 replicates.

Steps (SAS codes unless specified)

(1) Generate vector $\mathbf{f}^{(1)}$ = normal (repeat(**257715**,**24**))+ 10 mean values;/* mean values chosen randomly from a ranges such as 100 - 130, 50 - 58, etc.*/

(2) Generate V_c or V_r (say V_i) a correlation matrix for column or row observations. h = 0,1,2,3,4,5,6,7,8,9; r = 0.3; R = r^h; /* r stands for ρ */ print R; Construct a 10 X 10 correlation matrix as

V_i =

```
{1 0.3    0.09  0.027 0.0081   0    0    0         0      0,
0.3  1    0.3   0.09  0.027 0.0081       0    0    0      0,
0.09 0.3  1     0.3   0.09  0.027 0.0081       0    0      0,
0.027 0.09 0.3  1     0.3   0.09  0.027 0.0081      0      0,
0.0081    0.027 0.09  0.3   1     0.3   0.09  0.027 0.0081 0,
0    0.0081    0.027 0.09  0.3   1     0.3   0.09  0.027 0.0081,
0    0    0.0081    0.027 0.09  0.3   1     0.3   0.09  0.027,
0    0    0    0.0081    0.027 0.09  0.3   1     0.3   0.09,
0    0    0    0    0.0081    0.027 0.09  0.3   1     0.3,
0    0    0    0    0    0.0081    0.027 0.09  0.3   1};
```

(3) Generate correlated errors by defining

x = normal (repeat(**257715**,**24**));
H = root (V_i); L = T(H); /*transpose of H*/
e = L*__x__; /* construct error vector */

(4) Form correlated observations. **y**(1) = $\mathbf{f}^{(1)}$ + **e**; /* The sample **y**(1) will represent observations for replicate 1. Replicate 2 and 3 are simulated similarly except the randomization seed for generating **x** is changed. Form a complete sample (**y**) by appending Form a complete sample by appending **y**(1), **y**(2) and **y**(3). Append y(1), y(2), y(3);

(5) Generate a design in GenStat and display it in spread sheet. Define column positions as corresponding to 10 levels of crop B and rows as corresponding to 10 levels of crop A. Enter the row positions in ascending order.

AGHIERARCHICAL [Print=Design;Seed=32154]
BLOCKFACTORS=Block,Plots; TREATMENTFACTORS=*,!p(Treat2,Treat3);
LEVELS=3,!p(10,10) *3 blocks/replicates, 10 levels of both crops A and B*

Add the data into the spread sheet columns. This data is correlated only along rows, to include correlation along columns sort the spread sheet according to column (factor not to be confused with spread sheet column) levels in ascending order. Then add to the **y** column the replicated vector **e**.

(6) To add global trend along rows use T = 1.85 + 0.3 R_i - 0.15R_i^2. where average position number is 45/10 = 4.5. GenStat codes:

For n = 1 to 3
Calculate R = Row – 4.5
& T = 1.85 + 0.3*R – 0.15R**2
endfor

Add the vector **T** to column of **y**. Note the row positions should be arranged in ascending order.

Appendix C

TABLE V: Some average variances of treatment contrasts from a model with and without block or main plot effect term.

Model with block/main plot effects included	Model with block/main plot effects excluded
0.6638	1.0840
0.6605	0.7042
0.7011	2.0184
0.5766	2.2753
0.6605	0.6627
0.7214	0.8561
0.6846	3.5224
0.6669	2.3221
0.6270	2.2025

APPENDIX D

Simsim-finger millet data for Chapter 2 (Source: Makerere University Agricultural research Institute, Kabanyolo, Uganda)

Rep!	Row!	Splt!	SpFM	SpSM	YFM	YSM	RLERfm	RLERsm	RLERt	ER	LER_
1	1	1	60000	44400	518.7	503.3	0.6275	0.642	995.2	812.7	829.562
1	1	2	60000	22200	545.7	238	0.5534	0.3036	671.9	510.9	581.245
1	1	3	60000	11100	323.4	293.3	0.328	0.3741	550.4	455	496.719
1	1	4	60000	6330	397.7	250	0.0991	0.3189	327.7	298.9	500.153
1	2	4	40000	44400	455.97	313.3	0.6652	0.3996	834.8	641.2	600.105
1	2	3	40000	22200	485.2	277.8	0.4921	0.3543	663.6	520.4	582.991
1	2	2	40000	11100	349.2	213.3	0.3542	0.2721	491	387.9	432.947
1	2	1	40000	6330	342.2	173.3	0.3471	0.221	445.4	344.4	388.544
1	3	1	20000	44400	475.7	351.1	0.4825	0.4478	729.3	589	650.315
1	3	2	20000	22200	374.4	304.4	0.3797	0.3883	602.1	491.6	539.898
1	3	3	20000	11100	360.4	297.8	0.3655	0.3798	584.4	478	524.492
1	3	4	20000	6330	298.7	242.2	0.3029	0.3089	479.7	391.5	430.082
1	4	4	10000	44400	365.4	386.7	0.6748	0.4932	915.8	719.4	616.537
1	4	3	10000	22200	603.1	316.7	0.6117	0.404	796.2	618.2	696.05
1	4	2	10000	11100	377.4	290	0.3828	0.3699	590.1	478.7	527.385
1	4	1	10000	6330	363.4	286.7	0.3686	0.3657	575.7	468.4	515.279
1	5	1	0	44400	0	616.7	0	0.7866	616.7	616.7	*
1	5	2	0	22200	0	366.7	0	0.4677	366.7	366.7	*
1	5	3	0	11100	0	330	0	0.4209	330	330	*
1	5	4	0	6330	0	216.7	0	0.2764	216.7	216.7	*
2	5	1	60000	44400	401.11	383.3	0.6096	0.4889	861.3	683.9	635.598
2	5	2	60000	22200	407.4	350	0.4132	0.4464	673.9	553.7	606.255
2	5	3	60000	11100	364.7	286.7	0.3699	0.3657	576.7	469	516.096
2	5	4	60000	6330	344.7	266.7	0.3496	0.3402	540.8	439	483.516
2	4	4	40000	44400	571.7	266.7	0.6812	0.3402	800.8	602.5	626.299
2	4	3	40000	22200	401.7	173.3	0.4074	0.221	492.7	374.1	425.969
2	4	2	40000	11100	369.5	148.9	0.3747	0.1899	442.7	333.6	381.316
2	4	1	40000	6330	230.7	128.9	0.234	0.1644	312.3	244.2	274.01
2	3	1	20000	44400	309.8	493.3	0.9423	0.6292	1232.1	957.9	688.164
2	3	2	20000	22200	490.4	477.8	0.5988	0.6094	947.2	773	786.262
2	3	3	20000	11100	285.4	420	0.2895	0.5357	646.9	562.7	599.517
2	3	4	20000	6330	169.1	388.9	0.1715	0.496	523.4	473.4	495.264
2	2	4	10000	44400	308.2	196.7	0.7638	0.2509	795.5	573.2	390.558
2	2	3	10000	22200	627.3	193.3	0.7377	0.2466	771.7	557	587.872
2	2	2	10000	11100	597.4	193.3	0.6059	0.2466	668.3	492	569.065
2	2	1	10000	6330	302.1	156.7	0.5397	0.1999	579.8	422.8	346.721
2	1	1	0	44400	0	750	0	0.9566	750	750	*
2	1	2	0	22200	0	463.3	0	0.5909	463.3	463.3	*
2	1	3	0	11100	0	420	0	0.5357	420	420	*
2	1	4	0	6330	0	406.7	0	0.5187	406.7	406.7	*
3	1	1	60000	44400	427.4	373.3	0.4335	0.4761	713.1	587	742.135
3	1	2	60000	22200	412.4	270	0.4183	0.3444	597.9	476.2	529.4
3	1	3	60000	11100	381.7	180	0.3871	0.2296	483.5	370.9	420.089
3	1	4	60000	6330	187.4	123.3	0.1901	0.1573	272.3	217	241.175
3	2	4	40000	44400	616.4	268.9	0.6252	0.343	759	577.1	756.616
3	2	3	40000	22200	300.4	264.4	0.6089	0.3372	741.8	564.6	453.352
3	2	2	40000	11100	346.2	237.8	0.3511	0.3033	513.1	410.9	455.56
3	2	1	40000	6330	159.3	184.4	0.06	0.2352	231.5	214	284.6
3	3	1	20000	44400	453.2	635.6	0.8655	0.8107	1314.2	1062.3	920.663
3	3	2	20000	22200	230.1	380	0.7405	0.4847	960.5	745	524.733
3	3	3	20000	11100	633.2	231.1	0.7316	0.2948	804.7	591.8	629.383
3	3	4	20000	6330	255.1	148.9	0.2587	0.1899	351.7	276.4	309.358
3	4	4	10000	44400	696.97	333.3	0.808	0.4251	966.8	731.7	771.694
3	4	3	10000	22200	223.7	270	0.734	0.3444	845.4	631.9	410.707

3	4	2	10000	11100	608.7	266.7	0.5078	0.3402	664.8	517	649.572
3	4	1	10000	6330	392.1	203.3	0.3977	0.2593	515.1	399.4	449.931
3	5	1	0	44400	0	803.3	0	1.0246	803.3	803.3	*
3	5	2	0	22200	0	636.7	0	0.8121	636.7	636.7	*
3	5	3	0	11100	0	600	0	0.7653	600	600	*
3	5	4	0	6330	0	433.3	0	0.5527	433.3	433.3	*

Cowpea-sorghum data for Chapter 4 (Source: Serere Agricultural Research Institute, Uganda.)

PlotNo	Block	WP	S	sA	dC1!	dS1!	dC	dS	YC	YS	PredYC
1	1	2	1	1	8	105	224	1.0598	3.539	1.03774	
2	1	2	2	1	7	105	195	1.1628	2.84933	1.04245	
3	1	2	3	1	6	105	130	1.10412	2.90042	1.04715	
4	1	2	4	1	5	105	113	1.17704	2.62756	1.05186	
5	1	2	5	1	4	105	98	1.06055	2.73839	1.05657	
6	1	2	6	1	3	105	85.1	0.975447	2.88234	1.06128	
7	1	2	7	1	2	105	74	1.15407	3.36346	1.06599	
8	1	2	8	1	1	105	63.3	1.21722	3.17555	1.0707	
9	2	1	1	1	1	105	63.3	1.255	5.642	1.04464	
10	2	1	2	1	2	105	74	1.20416	5.28099	1.03993	
11	2	1	3	1	3	105	85.1	0.85262	5.11443	1.03522	
12	2	1	4	1	4	105	98	0.938201	4.17128	1.03051	
13	2	1	5	1	5	105	113	1.12401	5.17757	1.0258	
14	2	1	6	1	6	105	130	1.1445	5.67644	1.02109	
15	2	1	7	1	7	105	195	1.03821	5.25471	1.01638	
16	2	1	8	1	8	105	224	1.06911	5.71132	1.01167	
17	3	4	1	1	8	191	224	1.022	5.21	1.03778	
18	3	4	2	1	7	191	195	1.0626	4.90853	1.04249	
19	3	4	3	1	6	191	130	0.773909	5.14726	1.0472	
20	3	4	4	1	5	191	113	0.983474	4.28313	1.05191	
21	3	4	5	1	4	191	98	1.01373	4.78399	1.05661	
22	3	4	6	1	3	191	85.1	0.849917	5.35983	1.06132	
23	3	4	7	1	2	191	74	0.784873	5.02623	1.06603	
24	3	4	8	1	1	191	63.3	0.975465	5.48077	1.07074	
25	1	1	1	2	1	105	63.3	1.251	3.791	1.11923	
26	1	1	2	2	1	105	74	1.14361	3.21838	1.09581	
27	1	1	3	2	1	105	85.1	1.00405	4.28715	1.07238	
28	1	1	4	2	1	105	98	1.07457	4.29585	1.04896	
29	1	1	5	2	1	105	113	1.21652	4.1181	1.02554	
30	1	1	6	2	1	105	130	1.32645	4.06662	1.00212	
31	1	1	7	2	1	105	195	0.629904	4.24682	0.97869	
32	1	1	8	2	1	105	224	0.646706	3.50669	0.95527	
33	2	4	1	2	1	191	224	1.135	5.118	0.92921	
34	2	4	2	2	1	191	195	1.02021	5.20628	0.95263	
35	2	4	3	2	1	191	130	1.07788	5.25722	0.97605	
36	2	4	4	2	1	191	113	0.689641	4.92322	0.99947	
37	2	4	5	2	1	191	98	0.846725	5.87214	1.0229	
38	2	4	6	2	1	191	85.1	0.836116	5.89281	1.04632	
39	2	4	7	2	1	191	74	1.00102	5.12109	1.06974	
40	2	4	8	2	1	191	63.3	1.10149	5.7932	1.09316	
41	3	1	1	2	1	105	63.3	1.251	4.768	1.11927	
42	3	1	2	2	1	105	74	1.14471	4.20722	1.09585	

43	3	1	3	2	1	3	105	85.1	0.94779	5.37517	1.07242
44	3	1	4	2	1	4	105	98	0.883608	4.53206	1.049
45	3	1	5	2	1	5	105	113	1.06881	3.7281	1.02558
46	3	1	6	2	1	6	105	130	1.24159	4.86421	1.00216
47	3	1	7	2	1	7	105	195	0.750507	5.20477	0.97874
48	3	1	8	2	1	8	105	224	1.12314	4.63107	0.95531
49	1	3	1	1	2	1	191	63.3	1.163	4.259	1.03816
50	1	3	2	1	2	2	191	74	1.13209	4.00786	1.01918
51	1	3	3	1	2	3	191	85.1	0.83856	4.13118	1.00019
52	1	3	4	1	2	4	191	98	0.939675	3.65629	0.9812
53	1	3	5	1	2	5	191	113	1.18956	4.63091	0.96222
54	1	3	6	1	2	6	191	130	1.06204	4.99428	0.94323
55	1	3	7	1	2	7	191	195	1.0654	4.81034	0.92424
56	1	3	8	1	2	8	191	224	0.713388	4.24307	0.90526
57	2	2	1	1	2	8	105	224	0.966	5.272	0.87919
58	2	2	2	1	2	7	105	195	1.14702	4.26821	0.89818
59	2	2	3	1	2	6	105	130	0.795533	4.18	0.91717
60	2	2	4	1	2	5	105	113	0.698338	4.07752	0.93615
61	2	2	5	1	2	4	105	98	0.88019	4.83948	0.95514
62	2	2	6	1	2	3	105	85.1	1.01169	5.03044	0.97412
63	2	2	7	1	2	2	105	74	1.19193	5.43747	0.99311
64	2	2	8	1	2	1	105	63.3	0.742634	4.97866	1.0121
65	3	3	1	1	2	1	191	63.3	0.981	4.272	1.0382
66	3	3	2	1	2	2	191	74	1.25266	4.04436	1.01922
67	3	3	3	1	2	3	191	85.1	0.997562	4.33128	1.00023
68	3	3	4	1	2	4	191	98	0.970937	4.36354	0.98124
69	3	3	5	1	2	5	191	113	1.0832	4.07645	0.96226
70	3	3	6	1	2	6	191	130	0.993411	4.91885	0.94327
71	3	3	7	1	2	7	191	195	0.852378	4.81744	0.92429
72	3	3	8	1	2	8	191	224	0.803848	5.22578	0.9053
73	1	4	1	2	2	8	191	224	1.02	4.3	0.92001
74	1	4	2	2	2	7	191	195	0.876089	4.50587	0.93524
75	1	4	3	2	2	6	191	130	0.620916	4.74349	0.95047
76	1	4	4	2	2	5	191	113	0.894019	3.95248	0.96571
77	1	4	5	2	2	4	191	98	0.986074	3.7314	0.98094
78	1	4	6	2	2	3	191	85.1	0.904726	4.21139	0.99617
79	1	4	7	2	2	2	191	74	1.10048	4.39868	1.01141
80	1	4	8	2	2	1	191	63.3	1.20935	3.8144	1.02664
81	2	3	1	2	2	1	191	63.3	0.689	5.921	1.00057
82	2	3	2	2	2	2	191	74	0.622043	6.01907	0.98534
83	2	3	3	2	2	3	191	85.1	0.755548	5.71257	0.97011
84	2	3	4	2	2	4	191	98	0.787971	5.19014	0.95487
85	2	3	5	2	2	5	191	113	0.895531	5.86722	0.93964
86	2	3	6	2	2	6	191	130	1.00113	5.29755	0.92441
87	2	3	7	2	2	7	191	195	0.997394	5.45649	0.90918
88	2	3	8	2	2	8	191	224	0.754516	5.0821	0.89394
89	3	2	1	2	2	8	105	224	1.06	4.21	0.92005
90	3	2	2	2	2	7	105	195	1.12394	4.40221	0.93528
91	3	2	3	2	2	6	105	130	0.898669	4.6419	0.95051
92	3	2	4	2	2	5	105	113	1.04255	3.86108	0.96575
93	3	2	5	2	2	4	105	98	1.22494	3.68152	0.98098
94	3	2	6	2	2	3	105	85.1	1.10805	4.15105	0.99621
95	3	2	7	2	2	2	105	74	1.25242	4.16118	1.01145
96	3	2	8	2	2	1	105	63.3	1.28567	3.95564	1.02668

APPENDIX E

'Code for testing for error symmetry: SAS program'

```
data x;
input bloc Row Sbplt SA dA dB R1 R2 Y1; cards;
 proc mixed;
   class bloc Row Sbplt SA dA dB;
   model Y1 = dA|dB|SA/out=pred;
random bloc bloc*dB;
   repeated/ type = sp(spherical) (Row Sbplt) /* AR(1) also used */
             subject = bloc*SA*dA;

   ods output pred = C;
run;
 C = Y1 - C;
proc print data = C; run;
```

```
/*Then genstat6 was used to generate summary statistics */

DESCRIBE [SELECT=nobs,nmv,mean,median,min,max,q1,q3,var,kurtosis] C
```

'SAS Code for data handling and efficiency computations'

```
data x;
input bloc Row Sbplt SA dA dB R1 R2 Y1;/* Row = main plot, dA = crop A
density levels,
dB = crop B density leves, R1 = Sbplt position number - mean Sbplt
number, R2 = R1*R1 (second order polynomial; higher orders can
defined similarly, Y1 = yield */
cards;
;

proc mixed MMEq MMEqSol AsyCov; /* requests for mixed model equations
and solutions and asymptotic var-cov matrix of covariance parameters*/
   class bloc Row Sbplt SA dA dB;
   model Y1 = dA|dB|SA/ InvCovB; /* (1) Trend in Sbplt may be added by
introducing terms
   R1 or R2 or higher polynomial terms (2)InvCovB requests for inverse
of var-cov matrix of fixed effects */
   random bloc bloc*dB;
   random dB/ type = sp() (Row Sbplt) /* Define covariance model e.g.
sp(exp) for exponential */
             subject = bloc*SA*dA; /* This whole term is included if
there is no spatial dependence in data */
       /* or define repeated term as */
   repeated/ type = ar(1) /* arma(1,1) can be used instead of ar(1) */
             subject = bloc*SA*dA;
   lsmeans dA|SA/pdiff;
```

```
   estimate '' ; /* Define matrix for treatment combinations/contrasts
*/
   contrast '' ; /* Define contrast matrix */
   contrast ''; /* Define matrix for orders of orthorgonal polynomials
in dB */
   ods output contrasts = a;
run;

data power; set a;/* calculation of power of hypothesis testing under
contrasts */
  alpha =0.05;
  nc=numdf*Fvalue;
  fcrit = finv(1-alpha,numdf,dendf,0);
  power = 1 - probf(fcrit,numdf,dendf,nc);
proc print;
proc chart data=lsdiffs1;
   by effect notsorted;
   vbar stderr;
   run;

QUIT;


'Genstat Code for data handling - smoothing splines and AR(1)xAR(1)'
Vcom[Fixed = dA*SA*dB]random = bloc +Row.Sbplt;Cos=pos
Vstructure[Term = Row.Sbplt]AR,AR;Factor = Row, Sbplt;Initial
=!(0.6)!(0.6)
Reml[.....]Y1 "With splines the code is"

Vcom[Fixed = dA*SA*dB; Spline =Sbplt]random = bloc +Row.Sbplt;Cos=pos
Vstructure[Term = Row.Sbplt]AR,AR;Factor = Row, Sbplt;Initial
=!(0.6)!(0.6)
Reml[.....]Y1.
```