APPLICATION OF SPATIAL STATISTICAL METHODS TO THE STUDY OF SOIL FERTILITY PATTERN

BY

KINFEMICHAEL ALEMU

የአዲስ አበባ ዩኒቨርሲቲ
ቤተ መጻሕፍት
ADDIS ABABA UNIVERSITY
LIBRARIES

A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE

STUDIES OF ADDIS ABABA UNIVERSITY

IN PARTIAL FULLFILLEMENT OF THE REQUIREMENTS FOR

THE DEGREE OF MASTER OF SCINENCE IN STATISTISCS

JUNE, 2000

ADDIS ABABA

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

APPLICATION OF SPATIAL STATISTICAL METHODS TO THE STUDY OF SOIL FERTILITY PATTERN

BY

KINFEMICHAEL ALEMU

DEPARTMENT OF STATISTICS

FACULITY OF SCIENCE

APPROVED BY

DR. BUTTE GOTU
ADVISOR

DR. ESHETU WENCHEKO
EXAMINER

DR. TADEWOS KOROTO
EXAMINER

Acknowledgments.

It is a pleasure to acknowledge those who have assisted me at different stages in the preparation of this thesis. My first thanks are due to Dr. Butte Gotu, my advisor, for initially encouraging me to undertake the subject and spent his invaluable time by guiding me with detailed and thoughtful comments. My thanks are also to Ato Girma Taye, Head of Biometrics and Data Processing Services of The Ethiopian Agricultural Research Organization (EARO) for providing me with the data for this paper and sharing his technical expertise in agricultural experiments.

I am highly indebted to Ato Debebe Ero who provided me with a laptop so that I could work all the bulky and time consuming data analysis, writing, and editing of the paper at home. My special thanks are also to w/t Senaite G/Selassie who helpfully took the printing burdens for me.

I am also grateful to my colleagues at the I. C. S. Computer Training Center especially to w/t Rahel Tafesse for her unrestricted encouragement from the start and always helpful in arranging work schedules that comply with my regular class schedules which helped me to do both simultaneously.

Last but by no means least, my heart-full thanks go to my family at large and my sister, Mekdes especially, for being always there for me from the embryo stage to my eventual success.

Others still I have to thank but without mentioning their names since they are many and I wouldn't be able to do it in only a few pages. Thank you all.

ABSTRACT

Adequate characterization of soil heterogeneity of an experimental site is a good guide and at times even a prerequisite to choose a good experimentation technique. Based on data from a uniformity trial on wheat, soil fertility pattern of an experimental farm was studied using the techniques of uniformity trials and spatial autocorrelation analysis.

From the soil productivity contour map, a uni-directional fertility gradient (from left to right) was observed on fields using broadcast sowing method than using row plantation method. The same result was also confirmed using the mean square between stripes method. The distribution of fertile plots was found to be relatively heterogeneous for row planted fields than for broadcast sown fields. The degree of adjacent plot correlation between yield of plots using broadcast sowing method was relatively higher than for row plantation method as measured by Smith's coefficient of heterogeneity indicating that the variability in yield due to any given plot size does also depend on crop geometry (type of plantation method). Both square and rectangular plots were equally efficient in reducing the error variance. A plot size of 8 to 12 m^2 was found to be optimum on both plantation methods.

The Moran's, I, and Geary's, c, coefficients were used to test against spatial autocorrelation between yields of neighboring experimental plots. For all the three cases considered to represent 'neighborhood' between plots, namely, Bishop's case, Rook's case and Queen's case, both tests confirmed that there is a significant spatial

autocorrelation between neighboring plots using broadcast sowing method. For row planted fields, the observed spatial autocorrelation became insignificant for a plot size 30 m^2 with bishop's case, 4 m^2 with rook's case (using Moran's I), and 12 m^2 with queens case (using Moran's I).

REFERENCE

CHAPTER ONE **1**

INTRODUCTION **1**

1.1 *General* 1

1.2 *Background* 3

1.3 Spatial analysis 4

1.4 *Literature review* 8

1.4.1 Uniformity trials 8

1.4.2 Spatial analysis and autocrrelation 11

1.5 *Objectives of the study* 15

CHAPTER TWO **16**

DATA AND METHODOLOGY **16**

2.1 *The Data* 16

2.2 *Statistical Methods* 17

2.2.1 Uniformity trails 17

2.2.1.1 *Soil productivity countour map* 18

2.2.1.2 *Mean square between stripes* 20

2.2.1.3 *Smith's index of soil hetrogenity* 21

2.2.2 Spatial autocorrelation 27

2.2.2.1 The spatial autocorrelation problem 26

2.2.2.2 Tests againest spatial autocorrelation 27

CHAPTER THREE **35**

RESULT AND DISCUSSION **35**

3.1 *Results on uniformity trials* 35

3.1.1 Soil productivity Contour Map *35*

3.1.2 Mean square between stripes *39*

3.1.3 Smith's Index of soil heterogeneity *40*

3.1.4 Plot shape *49*

3.1.5 Plot Size *50*

3.2 *Spatial autocorrelation* 54

3.3 *Conclusion and recommendation* 67

APPENDIX 1 ..... **69**

MODELS FOR SPATIAL DATA **69**

A. *Why spatial models?* 69

B. *The general spatial model* 70

C. *Nearest-neighbor system on a lattice* 72

APPENDIX 2 ..... **75**

APPENDIX 3 ..... **77**

REFERENCES **79**

CHAPTER ONE

INTRODUCTION

1.1 General

The need to develop statistical theory for designing experiments stems, like the need for statistical analysis of numeric information, from the inherent variability of the experimental results. In the physical sciences this variability is relatively small and, when we think about experiments in physics and chemistry we usually think of 'correct results' from the experiment. However, even in these deterministic mathematical models, practical experiences show that the results are to some extent variable, mainly due to the complexities of the measurement process or from the inherent variability of the experimental material.

Unlike the physical sciences, experiments of the biological sciences are not exact. Most experiments on plants or animals use many plants or animals because it is clear that two plants or animals differ. In one way or the other, as the complexities of the experiments increase or the variation between the experimental units becomes large, and when the difference of interests is relatively small, then the precision of the experiment becomes more important.

The experimental procedure, which we use to maintain the results of an experiment at a given level of precision, usually consists of four phases (Gomez and Gomez, 1984).

- Selecting the appropriate material to test
- Specifying the characteristic to measure
- Selecting a procedure to measure the characteristics
- Specify a procedure to determine whether the measurement made support the hypothesis or not.

The first two are relatively simple for the subject matter specialist than the latter two which heavily depend on the design of experiments having three essential components: Estimation of error, error control, and proper interpretation of the results.

Every experiment should be designed to have a measure of the experimental error. Replication is the usual way of designing an experiment to have a measure of the experimental error. However, with several units for each treatment, the ambiguity that occurs when each treatment has a single unit is not completely eliminated. One or more treatments may be 'lucky' in the selection of units to which they are to be applied.

Consider a breeding experiment of variety A and B replicated as (Gomez and Gomez, 1984).

A	B	A	B	A	B	A	B

If the area has a uni-directional fertility gradient, say increasing fertility level from left to right, variety B would be a handicap because it is always to the right side of A which is relatively less fertile. Hence replication alone may not give us the exact measure of the experimental error or generally a good design; the way we replicate the treatments in different plots, experimental settings, ... etc do also matter. Randomization is one of the methods used to control systematic variations.

From the above discussion it is clear that, replication and randomization do give a good measure of the experimental error. But sometimes when we have some idea about which experimental units are likely to behave similarly, we can use blocking methods to make the allocation more 'fair'.

Generally a good experimental design should take into account all possible ways of measuring and controlling the experimental error so as to detect the true difference between the treatments being studied as the two are inversely related.

1.2 Background

In agricultural field experiments, the experimental units are small area of lands or plots of different sizes and shapes depending on the type of crop and treatment to be applied. Naturally we might expect that plots that are close to each other would behave more similarly than those further apart (of course the experimenter might have some knowledge about his experimental units). Either from a specific study of the fertility gradient of the soil over the experimental farm or from a general visual

inspection of the farm, plots will be grouped together in sets such that plots within a group might be expected to behave similarly than those in the other group.

The pattern of variation on which blocking systems can be based are very numerous and always the knowledge of the experimenter about his/her experimental material must be the most important source of information in determining blocks of units. Cochran and Cox (1957), Ostle and Mensing (1975), Gomez and Gomez (1984), Mead (1988) discussed different blocking strategies in various disciplines and experimental settings.

The present study deals with applications of different methods to study soil fertility pattern of a certain experimental farm, which can be used to select an appropriate blocking strategy for agricultural experimenters on that area.

1.3 Spatial analysis

Spatial data (spatially arranged measurements) occur in wide variety of ranges. The origin of human life link studies of the evolution of galaxies, the structure of biological cells and settlement pattern in archeology. The estimation of rainfall, and of ore, of petroleum reserves is of prime importance. Foresters and agriculturists need to study plant competition and account soil variation in their study. Economists study the variation of prices of different commodities or inflation rates among neighboring countries. The geographical arrangement of activities in urban areas has received a great deal of attention from economists, geographers, city planners, and ecologists.

Generally in almost all disciplines we find data, which in one way or the other are spatially (or geographically) located.

Spatial statistics deals with the analysis of spatially arranged measurements and/or patterns and reducing such measurements into few, clear and useful summaries. As the scope of statistics goes beyond a simple reduction of data to useful summaries to comparison of these summaries with what might be expected from theories of how the pattern might have originated or developed (confirmatory data analysis), spatial data analysis does also study the behavior of the stochastic process generating these spatial patterns and attempts to model the situation.

Spatial data are of many types, but basically we can categorize them into three: -

1) **point patterns** arise when the important variable to be analyzed is the location of 'events'. That is, when the size of the objects of the measurable variate are so small compared with the distance between them that their size may be ignored. Usually counts (or points) of occurrences or incidences are considered. Examples of this type are location of birds' nests, locations of imperfections in metals, or rocks, galaxies, towns, earthquakes, ...etc

2) **Lattice data** these are data arranged in a rectangular grid. They usually arise from experiments either when a deliberate systematic sampling is done or from agricultural field trials in which a field has been divided into

rectangular pixels. Apart from the usual systematic sampling of regular or irregular lattices, remote sensing from satellites do also offer an efficient means of data gathering. For example, it allows information to be gathered rapidly on weather patterns, mineral distribution, and crop acreage without conducting lengthy and labor-intensive traditional surveys.

3) **Geostatistical data** are when both location and size of the measurable variate are considered. Geostatistics, as a hybrid discipline of mining engineering, geology, mathematics, and statistics, mainly studies spatial variability at both large and small scale, that is, it models both spatial trend and spatial correlation. One of the most important problems in Geostatistics is to predict the ore grade in a mining block from observed samples. The method has found favor among soil scientists who seek to map soil properties of a field from a small number of soil samples at known location throughout the field.

When time dimension is involved, we get, *Spatio- temporal data.* These are spatial data measured in different time periods. Examples of spatio-temporal data are earthquake occurrences in different regions at different times, occurrence of an epidemic in successive years or seasons in different regions, ...etc.

In all types of spatial data, similarity among observation from neighboring regions is expected. When observations of a given variable show location (spatial) dependence, then we talk about spatial correlation. Unlike serial correlation in a time

series data, where at a given point there is a clear distinction between past and future values (uni-dimensional dependence), dependence in spatially located data extend in all directions. A fertilizer applied at a given plot affects all neighboring plots.

Spatial dependencies among observations of a given variate /s do also have a meaningful effect on the well-developed statistical methods like regression and time series analysis. The observed spatial dependence in the residuals of a certain regression model could be due to omission of important explanatory variables, incorrect specification of the relationship, measurement error or due to the location of the observed variates in space (see Cliff and Ord, 1973, p. 87). When the errors are spatially correlated, the usual ordinary least squares (OLS) method is no more the best estimate of the regression coefficients. Generally, in spatial data analysis, we may be confronted with the following problems:

How can we detect and measure spatial dependencies? What is the implication of observing significant spatial dependencies in statistical analysis? Can we adjust our model to involve these observed dependencies? Can we have a stochastic model to represent the process generating the series? Why do we need to model and study these spatial patterns? Statistics for spatial data deals with these and other related topics.

This study attempts to detect and measure the intensity of spatial autocorrelation on a lattice data generated from agricultural field trials by applying spatial statistical methods,.

1.4 Literature review

From the bulk of literature, only those that are more related and relevant to this study and some selected applications of the methods used in this study in other disciplines are presented. It's the author's belief that further references cited in each of the presented literatures will help the interested reader for further understanding.

1.4.1 Uniformity trials

To test a hypothesis of equality of treatments with predetermined confidence, experimental error should be kept to the minimum. In agricultural field experimentation this can partly be achieved by using optimum plot sizes with appropriate shape. The research to determine optimum plot size and shape through the techniques of uniformity trials started long ago. A real break-through has, however, been attained after Smith (1938), who showed empirically that the logarithm of variance between plots of a given size is linearly related to the logarithm of the size of the plot, and developed a method of estimation of soil heterogeneity index.

Narayana and Ramanata (1983) used Smith's empirical variance relationship to measure soil heterogeneity along square and rectangular plots of different sizes and shapes considering the number of basic units across and along the crop rows. They fitted the same model to a uniformity trial data on castor (Ricinun communis) variety

'Aruna', having duration of 150 days, during the rainy seasons of 1979 and 1980. They found higher values of the estimated heterogeneity coefficient for the year 1980 than in 1979. In both seasons, the trend in the value of b for the three different shapes, $x_1<x_2$, $x_1=x_2$, and $x_1>x_2$, was the same, where x_1 and x_2 represent the number of basic units across and along the crop rows respectively. This shows that the shape of the plots was equally efficient in reducing the error variance in both seasons. Rectangular plots with more number of basic units across the row than along the column were found to be more efficient. Using the comparable variance method, a rectangular plot of 12 basic units (17.28 m^2) was found to be optimum in both seasons.

In another study, Narayana and Ramanata (1985) introduced a modified use of Smith's empirical variance law to measure soil heterogeneity in square and rectangular plots of different sizes and shapes. Using a uniformity trial data on 'HY 2 pigeonpea', having duration of about 150 days, they found that the heterogeneity coefficient due to rectangular plots was highly significant; indicating a significant reduction in the error variance due to rectangular plots. Using the methods by Hathway (1961) a plot size of around 23 m^2 appeared to be optimum.

In estimating the regression coefficient of the logarithm of variance (V_x) on the logarithm of plot size (x), Smith recommended that the variances of different sized plots should be weighted by their respective degrees of freedom (the degree of freedom being the number of combined plots minus one). But since variance estimates for different size of plots are built up from common components, they are

frequently highly correlated. So a simple weighing by degrees of freedom will not give efficient estimates. Hathway and Williams (1985) show how efficient estimates of the constants in this relationship may be determined and presented a general method of determining efficient linear estimates when the data are correlated and of unequal variability.

Mohammed (1987) determined optimum plot size and shape for wheat experimentation using 1152 basic unit plots for three seasons in Sudan. It was found that an average plot size of 13.6 (54 m^2), for unguarded plots and 11.1 (44.4 m^2) for guarded plots appeared to be optimum over the seasons. Plots of different shapes were not significantly different in minimizing the error variance but long narrow plots were less variable.

Girma and Amsale (1999) attempted to study soil heterogeneity and detrmined an optimum plot size and shape for wheat experimentation in one of the former Institute of Agricultural Research (IAR), research sites – Holleta, Ethiopia. In their study, they found that there is a relatively high degree of correlation among adjacent plots in the study area for fields using broad cast sowing method than for row planted ones. Using the methods of Hathway (1961), they found an optimum plot size of 12.43 m^2 and 7.76 m^2 for broadcast sowing and row plantation methods respectively. Plots of equal size but different shapes were equally efficient in reducing the error variance. Crop- geometry (type of plantation method) had a significant effect on increasing degree of adjacent plot correlation and hence variability in yields of plots of different plot sizes due to high competition effect for broadcast sowing method.

Katayal and Rajput (1978), Gupta et al. (1991), Ram et al. (1993), ... etc applied the techniques of uniformity trials and determined optimum plot size and shape for different crops in different soil conditions.

Fertility pattern is not only studied using a uniformity trial which mainly basis on the premises that, a single crop variety cultivated uniformly would give uniform yields, but also by using the spatial patterns of soil properties and nutrient concentrations (Soil Chemistry). This helps besides giving detailed information on the soil fertility patterns for choosing a better blocking strategy, to provide a site specific farming practices that match agricultural inputs with regional crop needs. Studies of such types have been discussed by some researchers. For example, Cahn et al. (1994) studied spatial variation of soil organic C (SOC), soil water content (SWC), NO3-N, PO4-P and K. The range of spatial correlation was determined from semivariance analysis of the data and found to vary among and within fertility parameters.

1.4.2 Spatial analysis and autocrrelation

Experimenting on optimum plot size and shape alone may not help to measure or control experimental error. Spatial statistical methods may sometimes be needed to take into account spatial dependencies among yields of neighboring plots in their model.

A paper by Besag and Kempton (1986) adds to the development of spatial statistics by different authors in the analysis of agricultural field experiments, principally to control local variation in soil fertility. The authors describe four distinct applications of the use of neighboring plot values in the analysis of field experiments and illustrate each with different experimental data.

To control local variation in soil fertility, Besag and Kempton used systematically placed check plots containing standard treatment or variety. The yields from these check plots are then used to calculate a local fertility index that can be used to adjust the yields of the experimental plots. Check plots are mainly used to improve the efficiency of selection in plant breeding trials, when large number of genetic lines are assessed in un-replicated plots. For replicated experiments, the authors used an improved method of Papadakis (1937) which does not require added check plots but uses neighboring experimental plots corrected for treatment effects as internal check for local control of soil fertility variation. The main aim has been to produce more accurate estimates of treatment contrasts by removing any local linear trends from the data, attributed to the pattern of soil fertility. The method is based on first differences of plot values. Apart from fertility variation among neighboring plots, another cause of association arises from their interdependence, most commonly thought interplot competition. Interplot competition between plants of different varieties (or plants receiving different treatments) can lead to serious bias in variety means. They also tried to show that competition effects will generally give rise to systematic bias which will persist over replicated sites and years. The paper

discusses the use of a simultaneous autoregressive formulation of competition model, where plot values are directly related to the values of neighboring plots.

The fact that neighboring plots tend to give similar results than those far apart gave the problem another dimension: spatial autocorrelation of errors in regression or time series analysis. Most papers on spatial autocorrelation in regression analysis begin directly by defining the problem as one of spatial dependencies without referring to the missing variable, improper structural form, or measurement error problem. They propose Generalized Least Square (GLS) or some variant as a computational method for 'solving' the spatial autocorrelation problem. But Miron (1984) provides a relatively non-technical introduction to spatial autocorrelation and its implication in regression analysis. In the paper, spatial autocorrelation is not treated as a problem to be solved in and of itself, rather it is portrayed as a manifestation of the missing variable, improper structural form, or measurement error problem. The approach taken here has three advantages: i) it makes clear why spatial autocorrelation exists ii) it suggests how it might be corrected and iii) it suggests that considering spatial autocorrrelation as merely a statistical nuisance to be corrected can be dangerous. Other causes can also lead to a bias in the estimated marginal effect of included independent variables. The problem here is that, computational approaches such as the GLS which can correct for spatial dependency problem cannot identify these potential bias. Miron discusses with examples different methods for identification and estimation of spatial autocorrelation intensity among regression residuals. It also presents a method to 'correct' spatial autocorrelation under the assumption that the structure of the autocorrelation is known a priori.

Many scholars, notably Moran (1950) and Geary (1954), developed different statistics to measure spatial autocorrelation in spatially/geographically located data. Different authors in varying disciplines have used these measures in their studies.

For example Sukal and Wortenberg (1983) have used the Moran's I and the Geary's c statistics to detect the assumption of spatial autocorrelation to gene frequency surfaces using the isolation- by- distance model as an example.

Bonham et al. (1995) applied the same measures to study spatial relationships of blue grama (Bouteloua gracillis) with other specific characteristics of mixed grass community in south Arizona.

Zhang et al. (1995) performed spatial autocorrelation analysis of trace element contents of soil in the Tianjon plain area of china using Moran's I statistic. Others like Chou (1992), Czaplewski et al. (1994), Li et al. (1996), ...etc used these and other measures in their researche to detect and measure spatial autocorrelation between observations of variates located in space.

In our country, although some efforts has been done to determine optimum plot size and shape for few crops in different locations of the IAR –research sites using the methods of uniformity trials, to my knowledge, no study has been conducted using spatial statistical methods to study soil fertility patterns in agricultural field experiments. The present study endeavors to bridge the existing gap.

1.5 Objectives of the study

In the light of the above discussion, the paper attempts,

- to study the fertility pattern of the experimental farm at the IAR –research site -Holleta.

- to detect and measure the intensity of spatial autocorrelation over the experimental farm.

- to discuss and compare the methods of uniformity trials used to measure degree of autocorrelation between adjacent plots with spatial statistical methods used to measure autocorrelation in the experimental lattice.

- to supplement the original study and its results using alternative methods and discussions.

CHAPTER TWO

DATA AND METHODOLOGY

2.1 The Data

The data were generated by the Ethiopian Agricultural Research Organization (EARO), biometrics and data processing services under the objective of determining optimum plot size, shape and number of replications for wheat experimentation in one of the former IAR research sites- Holleta.

In the experiment, wheat was grown under a uniform treatment, management and cultural practices on an area of 0.2 hectar for two consecutive seasons in 1996/1997 at Holleta research center farm. In both seasons, the experiment was done on the same farm but different locations. Twenty-centimeter spacing row plantation and broadcast sawing methods were used. To alleviate the problem of edge effects (see Dyke and Grundy, 1988, p. 21), at the time of harvest, guard rows were excluded and the remaining net area (area harvested), the area from which the crop was actually harvested, was divided into sub plots of one meter by one meter. Grains from each basic unit were harvested, bugged, threshed, cleaned, dried, weighed, and recorded separately. Sub plot yields will later be combined to simulate plots of different sizes and shapes in the study.

2.2 Statistical Methods

The methodological part of the study is broadly categorized into two parts. The first deals with characterizing the fertility pattern of the experimental farm and measures the correlation among adjacent plots using the various techniques of uniformity trials. The second attempts applying spatial statistical methods to test against spatial autocorrelation among 'neighboring' plots over the experimental farm.

2.2.1 Uniformity trails

Uniformity trials are used to study the fertility pattern of a certain experimental farm. The experimental farm will first be partitioned into small plots of land called basic units. A single crop variety will be planted and all cultural and management practices will be applied as uniform as possible for all the basic units within the experimental farm. All other sources of variation in yield like solar energy, plant population, pest incidence, ...etc are kept constant. Yields from each basic unit will be harvested and recorded separately. The difference in yield from each basic unit is then attributed to the naïve soil fertility variation of the farm. But, since difference in yield among basic units can also be due to various sources as stated above, the experimental site should be selected by taking care of the following conditions which aggravate soil fertility variation and hence difference in yield among basic units: Slopes, areas used for previous experimentation in different crops, graded areas, presence of large

trees, poles, or structures, and unproductive sites for treatment and usage (see Gomez and Gomez ,1984, p. 478).

Under the above premises, the following techniques will be applied:

2.2.1.1 Soil Productivity contour map

Soil productivity contour map is a simple but informative presentation of soil heterogeneity. The map describes graphically the productivity level of an experimental site based on moving averages of contiguous units.

The steps involved in constructing a fertility contour map are the following (Gomez and Gomez, 1984):

Step 1:- Describe the number of units to go into each moving average, say k. The area involved in each moving average should be as square as possible.

Step 2:- Compute the $\sqrt{k}$ by $\sqrt{k}$ moving averages corresponding to the s^{th} row and t^{th} column as:

$$P_{s,t} = \frac{\sum_{i=s-1}^{s+\sqrt{k}} \sum_{j=t-1}^{t-\sqrt{k}-1} Y_{ij}}{k}$$

where s = 2, ..., (r-1), and t = 2, ..., (c-1)

Y_{ij} represents the grain yield of the $(i,j)^{th}$ basic unit.

r = total number of rows, and c = total number of columns

The object of combining several basic units into single moving average is to reduce the large expected variability on small plots

Step 3:- Group the values of the moving averages into m classes. The classes are generated by dividing the ranges of the moving averages into m equal classes.

Step 4:- Assign different shading patterns to each of the m- classes. The darkest shade will be assigned to the class with the highest yield, the second darkest shade to the class with the second highest yield, ... ,the lightest shade to the class with lowest yield (helps for an easier visualization of the fertility pattern)

Step 5:- Draw the productivity contour map by applying the above shading patterns to the generated moving average lattice data.

Based on the resulting productivity contour map, the change in the level of fertility of the soil (fertility gradient) over the experimental farm can easily be visualized and blocking strategies chosen accordingly.

2.2.1.2 Mean Square Between Stripes

The procedure is generally used to test the randomness of the data set. However, it is also useful in the characterization of the trend in the underlined soil fertility using a uniformity trial data.

Instead of working with the original basic units, the units are first combined into sums of vertical and horizontal stripes. Variability between the stripes in each direction is then measured by the mean square between the stripes. The relative size of the two mean squares, one for the horizontal stripes and the other for the vertical stripes, indicates the possible direction of soil fertility gradient and suitable orientation of blocks.

The steps required are (Gomez and Gomez 1984):

Step 1:- compute the totals for the horizontal and vertical stripes as:

$V_i = \sum_{i=1}^{r} Y_{ij}$ = Total of the i^{th} vertical stripe

$H_j = \sum_{j=1}^{c} Y_{ij}$ = Total of the j^{th} horizonatl stripe

Step 2:- Compute the sum of squares between the horizontal and vertical stripes as:

$$VSS = \frac{\sum_{i=1}^{c} V_i^2}{r} - \frac{G^2}{rc}$$

$$HSS = \frac{\sum_{j=1}^{r} H_j^2}{c} - \frac{G^2}{rc}$$

where

$$G = \sum_{j=1}^{c}\sum_{i=1}^{r} Y_{ij}$$

Step 4:- Compute the mean square for the horizontal and vertical stripes as:-

$$VMS = \frac{VSS}{c-1}$$

$$HMS = \frac{HSS}{r-1}$$

Step 5:- Compare and discuss the resulting mean squares.

2.2.1.3 Smith's Index of Soil Heterogeneity

Coefficients of the previous types are usually used to study patterned field variabilities. But field variations can also be of non-patterned type where variations exist but not in a clear direction of fertility gradient.

Non- Patterned sources of variations are usually approached by the way of fairfield Smith (Smith, 1938). Smith's index of soil heterogeneity is used primarily

The index is calculated from the empirical relationship between plot variance and plot size given as:

$$V_x = \frac{V_1}{x^b} \qquad (1)$$

where V_1 = Variance between basic units

V_x = Variance between basic unit plots of size x m^2

x = plot size

b = Smith's index of soil heterogeneity

The value of the index lies between 0 and 1 and measures the correlation between adjacent plots in an experimental farm. A value close to 0 indicates perfect correlation and one close to 1 indicates a relatively low correlation. To study the degree of adjacent correlation and reduction in variance among yield of plots of different sizes and shapes, the following modified Smith's index can be used (Narayana and Ramanata, 1985):

$$V_x = \frac{V_1}{x_1^{b_1} x_2^{b_2}} \qquad (2)$$

which can be written as:

$$V_x = \begin{cases} \dfrac{V_1}{(x_1 x_2)^{B_1} (x_2 / x_1)^{B_2}} & \text{if } b_2 > b_1 \\ \dfrac{V_1}{(x_1 x_2)^{B_1} (x_1 / x_2)^{B_2}} & \text{if } b_2 < b_1 \end{cases} \qquad (3)$$

where b_1 and b_2 correspond to rectangular plots having longer side along and across the crop rows respectively. B_1 and B_2 are index of heterogeneity due to square and rectangular plots respectively. Girma and Amsale (1999) fitted the modified model as

$$B_1 = \frac{b_1 + b_2}{2} \qquad B_2 = \frac{|b_2 - b_1|}{2}$$

given in (2) and studied the reduction in variance due to rectangular plots to the same data. In this study we apply Smith's index as given in (1) to study the pattern of reduction in variance to both square and rectangular plots.

Steps for computing the index are the following:

Step 1:- Combine the r by c basic units to simulate plots of different sizes and shapes. Only combinations that fit exactly into the whole area (those combinations where the product of number of simulated plots and number of basic units per plot equal to the total number of basic units) are used.

Step 2:- For each of the simulated plots constructed in step1, compute the yield total, T, as the sum of the x basic units combined to construct that plot.

Step 3:- Compute the between plot variance $V_{(x)}$ as:

$$V_{(x)} = \frac{\sum_{i=1}^{m} (Ti - \overline{T})^2}{m - 1}$$

Where m is the total number of simulated plots of size x and $\bar{T}$ is the average yield per a plot size of x basic units.

Step 4:- For each plot size and shape, compute the variance per unit area as:

$$V_x = \frac{V_{(x)}}{x^2}$$

Step 5:- For each plot size having more than one shape, test the homogeneity of the between plot Variance ($V_{(x)}$) to see the significance of plot orientation (plot shape), on the fertility variation using the F-test or Chi-square test.

That is if there are only two shapes for a given plot size, use the F-test as:

$$F = \frac{V_{(x)}{}^1}{V_{(x)}{}^2}$$

where $V_{(x)}{}^1$ and $V_{(x)}{}^2$ are the between plot variances of the two plots with the same size.

Otherwise, if there are more than two shapes for a given plot size use the Bartlett's Chi-square test as cited in Gomez and Gomez (1984 p. 468):

$$\chi^2 = \frac{(2.3026)(f)(k \log V_{(x)p} - \sum_{i=1}^{k} \log V_{(x)})}{1 + [(k+1)/34f]}$$

where: k = number of different plot shapes of size x

f = number of basic units in each plot

$V_{(x)p}$ = (Pooled estimate of Variance) = $\frac{\sum_{i=1}^{k} V(x)}{k}$

Step 6:- For each plot shape whose plot–shape effect is non-significant, compute the average of V_x values over all plot shapes, for others use the lowest value.

Step 7:- Using the values of the variance per unit area, V_x, computed in the previous steps, estimate the regression coefficient bewteen V_x and plot size x. Smith's regression coefficient will be computed after linearizing the empirical low given in (1)

$$E(\log V_x) = E(\log V_1) - b \log x \qquad (4)$$

In estimating the coefficient b by β, the value of V_1 is assumed to be constant. Under this assumption, the least square estimate of b is

$$\beta = \frac{\log V_1 \sum \log x - \sum \log V_x \log x}{\sum (\log x)^2} \qquad (5)$$

2.2.2 Spatial autocorrelation

2.2.2.1 The Spatial autocorrelation problem

A basic property of spatially located data is that the set of values, $\{x_i\}$, are likely to be related over space. Many authors in various disciplines discuss this dependence among observations of variates in space. Some as cited in Cliff and Ord (1981) are summarized as follows.

Galton (1889) argued that the historical links between societies, through migration or cultural diffusion, induces dependence among observations. Galton heavily stresses that full information should be given as to the degree in which the customs of the tribes and races that are compared together are independent. It might be true, that some of the tribes had derived them from a common source, so that they were duplicate copies of the same original.

Tobler (1970) has referred to "the first law of geography: every thing is related to everything else but near things are more related than distant things."

And finally Heplle (1978) gives the following remark by the statistician, Stephan (1934), "data of geographic units are tied together like bunches of grapes, not separate like balls in urn."

If the $\{x_i\}$'s display interdependence over space, we say the data are spatially autocorrelated. A more formal definition may be given as follows (Cliff and Ord, 1973):

Let there be a study lattice which is exhaustively partitioned into n = rc non-overlapping counties, where r and c represent the total number of rows and columns respectively. For the sake of simplicity, we shall consider the experimental farm as the study lattice and 'plots' to refer counties (partitioned regions). Further, let X be a measurable variate, such that x_i is the observed value of X at the i^{th} plot. Now, If for every pair of counties, i and j in the study area the drawings which yield x_i and x_j are uncorreleted, then there is no spatial autocorrelation in the county system on X. Conversely, spatial autocorrelation exits if drawings are not all pairwise uncorreleted (or drawings are influenced by neighboring plots).

2.2.2.2 Tests against spatial autocorrelation

Various tests against spatial autocorrelation have been proposed by different authors in the literature according to the type of data used (nominal, ordinal, or interval scaled). In this study we use two widely applied test statistics to test for the significance of spatial autocorrelation.

Let $\{\delta_{ij}\}$ be a connection matrix used to represent 'neighbor hood' in the county system, in which $\delta_{ij} = 1$ if the i^{th} and j^{th} counties are 'joined', and $\delta_{ij} = 0$ otherwise, for i,j = 1,... n.

The first coefficient is according to Moran (1950) and is given by:

$$I = \frac{n}{2A} \frac{\sum_{2} \delta_{ij} z_i z_j}{\sum_{i=1}^{n} z_i^2},$$

and the second suggested by Geary (1954)

$$c = \frac{(n-1)\sum_{(2)} \delta_{ij} (x_i - x_j)^2}{4A \sum_{i=1}^{n} z_i^2}$$

where

$$\sum_{(2)} = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n}$$

$$z_i = x_i - \bar{x}$$

$$A = \frac{1}{2} \sum_{i=1}^{n} L_i$$

L_i = number of plots 'joined' to the i^{th} plot

Observe that both coefficients I and c take on the classical form of any autocorrelation coefficient in which the numerator is a measure of correlation among the {xi}'s and the denominator a measure of variance.

The connection matrix, $\{\delta_{ij}\}$, used to formalize the concept of a 'join' may take any of the following three forms:

I) $\delta_{ij} = 1$ if the i^{th} and j^{th} counties have a common edge, and $\delta_{ij} = 0$ otherwise. By analogy with chess moves, this is the so-called Rook's case.

II) $\delta_{ij} = 1$ if the i^{th} and j^{th} counties have common vertex, and $\delta_{ij} = 0$ otherwise; the Bishop's case.

III) $\delta_{ij} = 1$ if the i^{th} and j^{th} counties have a common vertex or edge, and $\delta_{ij} = 0$ otherwise; the Queen's case.

As discussed in chapter one, these coefficients can also be used to measure the degree of spatial autocorrelation in different types of spatial data than the regular lattice used in this study. In such conditions, instead of using binary weights, δ_{ij}, to operationalize the concept of a 'join', we may use a generalized weight matrix **W**, $\mathbf{W} = \{w_{ij}\}$, where w_{ij} denotes the weight or degree of association between counties i and j. The above coefficients can be generalized as:

$$I = \frac{n \sum_{(2)} w_{ij} z_i z_j}{S_o \sum_{i=1}^{n} z_i^2}$$

and

$$c = \left(\frac{n-1}{2 S_o} \right) \frac{\sum_{(2)} w_{ij} (x_i - x_j)^2}{\sum_{i=1}^{n} z_i^2}$$

where besides the previously defined notations,

$$S_o = \sum_{(2)} w_{ij}$$

Cliff and Ord (1973 ch.2) derive the moments of the coefficients I and c, and analyze their asymptotic properties under either of the two assumptions:

(1) Normality; here $\{x_i\}$'s are assumed to be the results of n independent drawings from a normal population or populations.

(2) Randomization; whatever the underlining distribution of the population or populations that generate the observations, we consider the calculated value of I or c relative to the set of all possible values which I or c could take on if {xi}'s were repeatedly permuted around the county system (there are n! such values).

Both coefficients statistics are asymptotically normally distributed under either of the normality or randomization assumptions for fairly general conditions (see Cliff and Ord, 1981, ch.2). Using the subscripts N and R to denote the normality and Randomization assumption respectively, we present the first two moments for both coefficients.

Let A = the total number of joins in the county system

$$D = \frac{1}{2}\sum_{i=1}^{n} L_i(L_i - 1)$$

$$S_1 = 4A(n^2 - 3n + 3) - 8(A + D)n + 12A^2$$

$$S_2 = 4A(n^2 - n) - 16(A + D)n + 24A^2$$

Then,

$$E_N(I) = E_R(I) = \frac{-1}{n-1}$$

$$E_N(I^2) = \frac{S_1 - 4A(3n-3)}{4A^2(n^2-1)}$$

$$Var_N(I) = \frac{(n-1)(S_1 - 12A(n-1) - 4A^2(n+1)}{4A^2(n^2-1)(n-1)}$$

$$E_R(I^2) = \frac{nS_1 - b_2S_2}{4A^2(n-1)^{(3)}}$$

$$Var_R(I) = \frac{(n-1)(nS_1 - b_2 S_2) - 4A^2(n-1)^{(3)}}{4A^2 n(n-1)^2(n-2)}$$

and

$$E_N(c) = E_R(c) = 1$$

$$Var_N(c) = \frac{(2A+D)(n-1) - 2A^2}{(n+1)A^2}$$

$$Var_R(c) = \frac{2A^2 S_3 + 2A(n-1)S_4 + (A+D)S_5}{n(n-1)(n-3)2A^2}$$

where

$$S_3 = -(n-1)^2 b_2 + (n^2 - 3)$$

$$S_4 = -(n-1)b_2 + n^2 - 3n + 3$$

$$S_5 = n^2 - n + 2)b_2 - (n^2 + 3n - 6)$$

$$n^{(b)} = n(n-1)...(n-b+1)$$

$b_2 = \frac{m^4}{m_2^2}$, is the sample Kurtosis coefficient, m_j being the j^{th} sample moment about the sample mean.

Under the null hypothesis of no spatial autocorrelation, the two coefficients can be used as a test statistics. For large n, the critical values can be obtained from standared normal table after standardization:

$$\frac{I - E(I)}{\sigma(I)}$$ For Moran's, I and

$$\frac{E(c) - c}{\sigma(c)}$$ For Greay's c statistics.

where $\sigma(I)$, $\sigma(c)$ are the standard devation of the I and c coefficients.

Geary's coefficient is structured so that under the null hypothesis of no spatial autocorrelation , c = 1, values of c which are less than 1, indicate positive spatial autocorrelation, and values greater than one indicate negative spatial autocorrelation. Adjustments in the standard normal deviate were made simply to correspond positive values with positive spatial autocorrelation.

Remarks:

1. From the definition of c and I coefficients and the notion of 'join' used in this study, it is clear that the statistics measure spatial autocorrelation between counties which are only first nearest neighbors (first order special autocorrelation). Thus, correlellogram analysis is required to see how the autocorrelation decays in space or 'joins' can be defined as 'second

nearest', 'third nearest' ...etc, to study the degree of autocorrelation between counties two or three distances apart.

2. A general stochastic model can be defined to represent first order spatial dependencies using different methods (see theoretical discussions given in Appendix 1 of this study).

CHAPTER THREE

RESULT AND DISCUSSION

3.1 Results on uniformity trials

Adequate characterization of soil heterogeneity of an experimental site is a good guide and at times even a prerequisite to choose a good experimentation technique. Based on the methods and procedures discussed in Chapter Two results of the survey with discussion will be presented in the order they appear in the previous chapter.

From 960 basic units for row planted and 720 basic units of broadcast sown (after guard rows excluded), the averages of yield per plot were 322.37 and 245.09 g/m^2 respectively. Though the experiment was conducted at a given experimental farm in the same season but in different locations, apart from the naïve soil fertility variations in the specific experimental sites, due to the uncontrolled plant population which will result in severe competition between adjacent plants for broadcast sown field, the observed difference in yield averages is as expected.

3.1.1 Soil productivity Contour Map

The productivity contour maps for both broadcast sown and row planted fields are given in Figures 1 and 2. The size of the moving average was choosen to be 3 by 3 meters. The resulting moving averages were grouped into six classes where the

classes are generated by dividing the 580 moving averages (with values from 131 to 368) for broadcast sown and 780 moving averages for row planted (with values from 196 to 410) into six equal intervals.

As can clearly be seen from Figure1 there is a uni-directional (left to right) fertility gradient with more fertile plots to be found in the right side of the experimental farm. The change in the productivity level is gradual than in patches.

For row planted plots, the distribution of fertile plots seems more 'fair' or 'even' than for broadcast sown. Adjacent plots produce more similarly than those far apart. Fertile plots are found relatively in central parts than in the left and right sides of the experimental farm. In situations where there is a clear fertility gradient, blocks can be chosen parallel to the fertility gradient but for experimental sites with no clear fertility gradient, an optimum plot size for experimentation and proper block size and shapes should be selected using different statistical methods given in the literature cited.

Figure 1 Soil Productivity contour Map of a field for broadcast sown measuring 10 X 58 basic units constructed from the moving averages of 3 X 3 basic units.

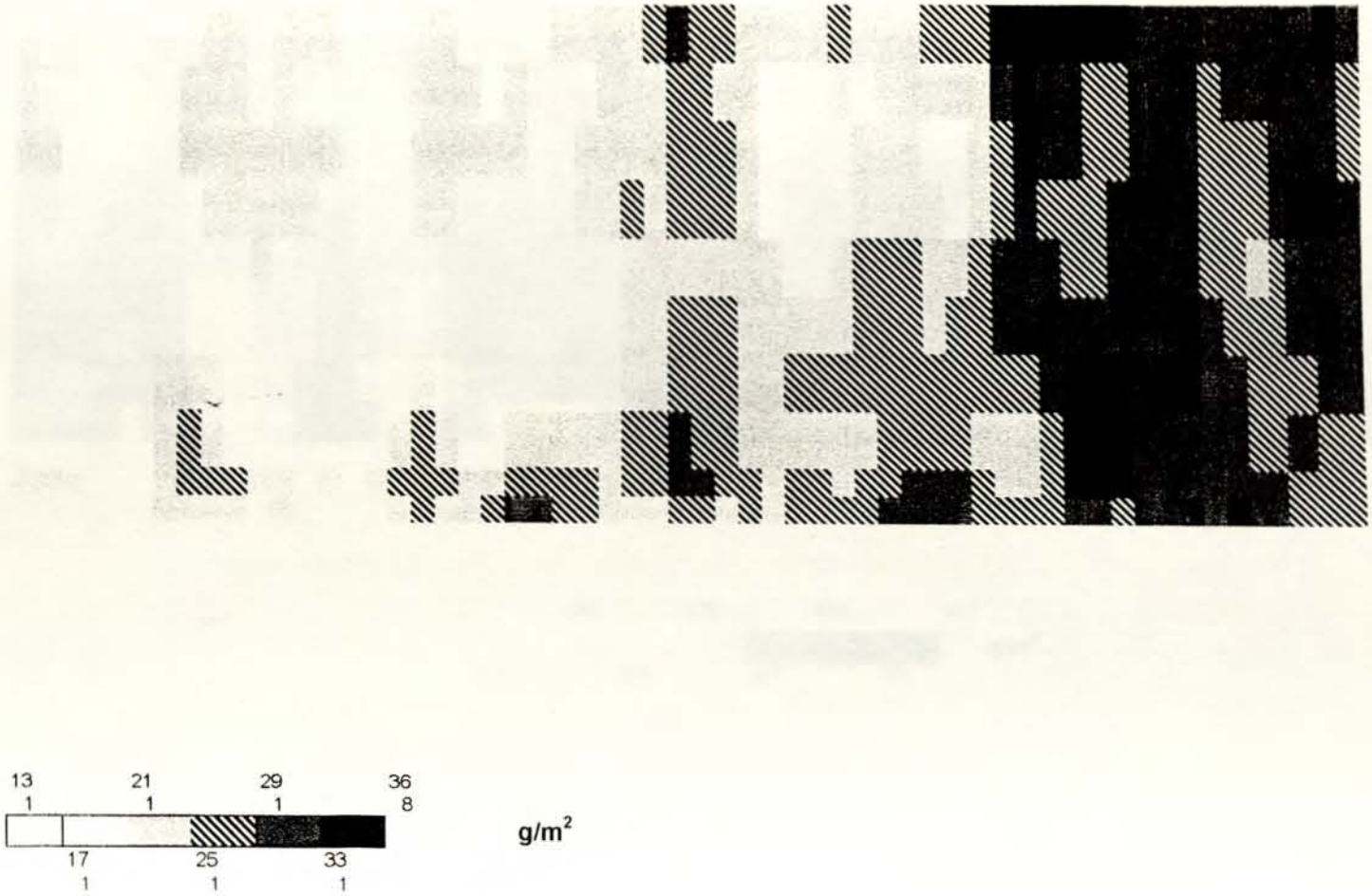
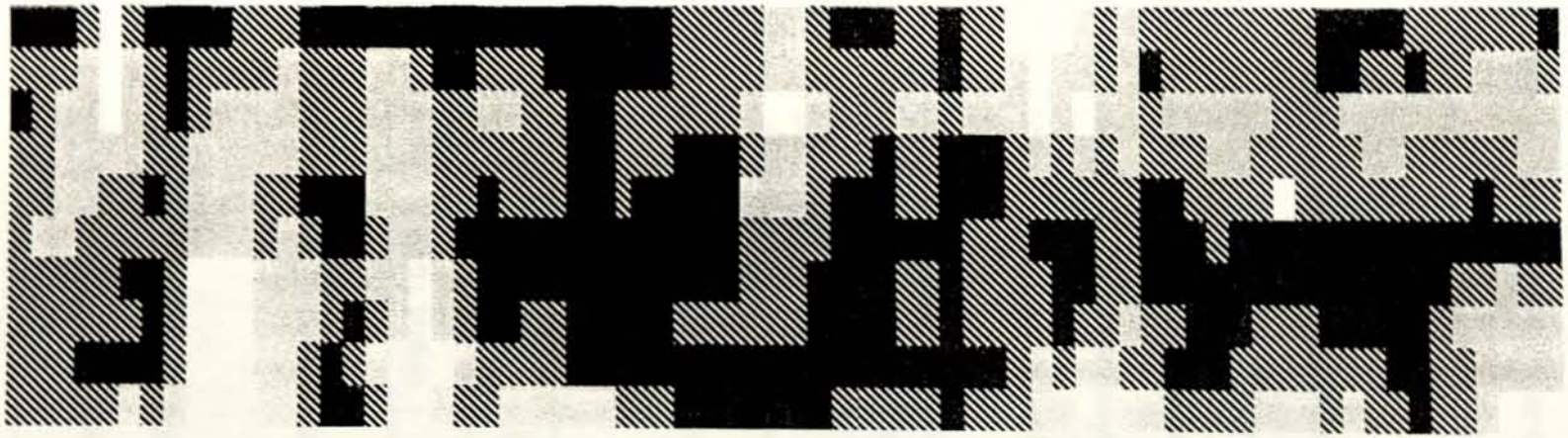


Figure 2 Soil Productivity contour map of a field for row planted measuring 10 X 78 basic units
Constructed from the moving averages of 3 X 3 basic units.



3.1.2 Mean square between stripes

Results on mean square between stripes for both plantation methods, broadcast sown and row planted are given in Tables 1 and 2.

Table 1: Summary of mean square between stripes for Broadcast sown

Source	r	c	Grand Total (G)	Sum of Squares	Mean Squares
Between Vertical stripes	12	60	176468	1823142.4	30900.7
Between Horizontal Stripes				332703.0	30245.7

Table 2: Summary of mean square between stripes for Row Planted

Source	r	c	Grand Total (G)	Sum of Squares	Mean Squares
Between Vertical stripes	12	80	309156	1121409.0	14195.1
Between Horizontal Stripes				422365.9	38396.9

From Table 1 we see that the Vertical–Stripes Mean square for broadcast sown is greater than the Horizontal–Stripes Mean square. This shows that the trend of soil fertility is more pronounced along the width than along the length of the field. This change in the fertility level was clearly seen as a uni-directional fertility gradient through visual inspection of the productivity contour map of Figure 1.

For row planted, however, the Between Horizontal–Stripes Mean square is more than twice the Vertical –Stripes Mean square indicating that the trend of soil fertility is higher along the length than the width of the field. Since, heterogeneity of plots is

more pronounced for row planted fields, the above justification may not be clearly seen from the productivity contour map of Figure 2.

3.1.3 Smith's Index of soil heterogeneity

Direction of fertility gradients were discussed using the previous two methods, namely the productivity contour map and mean square between stripes. To study the degree of correlation between adjacent experimental plots, Smith's index will be computed using the methods given in Chapter 2. All the 60 different sizes and shapes for broadcast sown and 54 for row planted were considered by combining basic units.

Tables 3 and 4 show results for both plantation methods. For plots of the same size but different shapes, tests for homogeneity of variance were done using the F-test and Bartlett's Chi-Square tests discussed in the methodology part and are presented in Tables 5 and 6.

None of the tests for homogeneity of variance considered for plots of the same size but different shape were significant under both plantation methods. This shows that both square and rectangular plots reduce error variance at the same rate. The average of $V_{(x)}$ values was considered in computing the coefficient b.

Table 3 Between-plot variance [V(x)], variance per unit area(Vx), and coefficient of variability (CV) for plots of various sizes and shapes, calculated from wheat uniformity data using broadcast sowing method.

Plot size and shape						
Size, m^2	Width, m	Length, m	Total no of plots	V(x)	Vx	CV
1	1	1	720	5884	5884	31.3%
2	1	2	360	19004	4751	28.1%
3	1	3	240	36835	4093	26.1%
4	1	4	180	60376	3773	25.1%
5	1	5	144	87245	3490	24.1%
6	1	6	120	119074	3308	23.5%
10	1	10	72	289482	2895	22.0%
12	1	12	60	412081	2862	21.8%
15	1	15	48	614519	2731	21.3%
20	1	20	36	1051833	2630	20.9%
30	1	30	24	2003641	2226	19.3%
2	2	1	360	17186	4297	26.8%
4	2	2	180	58565	3660	24.7%
6	2	3	120	116618	3239	23.2%
8	2	4	90	195104.0744	3049	22.5%
10	2	5	72	279734	2797	21.6%
12	2	6	60	391046	2716	21.3%
20	2	10	36	984702.4444	2462	20.3%
24	2	12	30	1379443.857	2395	20.0%
30	2	15	24	2118932.493	2354	19.8%
40	2	20	18	3748983	2343	19.8%
60	2	30	12	7189974.788	1997	18.2%
3	3	1	240	32784.31269	3643	24.6%
6	3	2	120	113329.6762	3148	22.9%
9	3	3	80	225456	2783	21.5%
12	3	4	60	384532.965	2670	21.1%
15	3	5	48	552857.0993	2457	20.2%
18	3	6	40	1309716.718	4042	26.0%
30	3	10	24	1975598.493	2195	19.1%
36	3	12	20	2734216.358	2110	18.7%
45	3	15	16	16406977.45	8102	36.7%
60	3	20	12	7656554.97	2127	18.8%
90	3	30	8	14584638.57	1801	17.3%
4	4	1	180	54644.97381	3415	23.9%
8	4	2	90	190303.9171	2973	22.3%
12	4	3	60	386725.1345	2686	21.2%
16	4	4	45	672022.9828	2625	20.9%
20	4	5	36	950317.6444	2376	19.9%
24	4	6	30	1362209.789	2365	19.8%
40	4	10	18	3547182.536	2217	19.2%

Table3 ctd.

Plot size and shape						
Size, m^2	Width, m	Length, m	Total no of plots	V(x)	Vx	CV
48	4	12	15	4971044.552	2158	19.0%
60	4	15	12	7927635.152	2202	19.2%
80	4	20	9	14277935.53	2231	19.3%
120	4	30	6	27948644.27	1941	18.0%
6	6	1	120	106068.0123	2946	22.2%
12	6	2	60	379131.4056	2633	20.9%
18	6	3	40	773988.8821	2389	19.9%
24	6	4	30	1345486.892	2336	19.7%
30	6	5	24	1954844.232	2172	19.0%
36	6	6	20	2822851.726	2178	19.0%
60	6	10	12	7310618.242	2031	18.4%
72	6	12	10	10362290.18	1999	18.2%
90	6	15	8	16556064.57	2044	18.5%
120	6	20	6	30769216.27	2137	18.9%
180	6	30	4	60747614	1875	17.7%
12	12	1	60	370808.626	2575	20.7%
24	12	2	30	1383762.271	2402	20.0%
36	12	3	20	5039121.368	3888	25.5%
48	12	4	14	2663991.258	1156	13.9%
60	12	5	12	7598650.788	2111	18.8%
72	12	6	10	11292492.62	2178	19.1%
120	12	10	6	31395921.87	2180	19.1%
144	12	12	5	45896591.3	2213	19.2%
180	12	15	4	75848054.67	2341	19.7%
240	12	20	3	151102836.3	2623	20.9%
360	12	30	2	359763488	2776	21.5%

Table 4 Between-plot variance [V(x)], Variance per unit Area(Vx), and Coefficient of variability (CV) for plots of various sizes and shapes, calculated from Wheat Uniformity data using Row plantation method.

Plot size and shape						
Size, m^2	Width, m	Length, m	Total no of plots	V(x)	Vx	CV
1	1	1	960	4563	4563.226	21.0%
2	1	2	480	13124	3281.001	17.8%
4	1	4	240	32356	2022.238	13.9%
5	1	5	192	50354	2014.159	13.9%
8	1	8	120	89363	1396.293	11.6%
10	1	10	96	140710	1407.098	11.6%
16	1	16	60	237487	927.6848	9.4%
20	1	20	48	362132	905.3294	9.3%
40	1	40	24	963494	602.184	7.6%
2	2	1	480	12256	3063.918	17.2%
4	2	2	240	35514	2219.611	14.6%
8	2	4	120	80222	1253.468	11.0%
10	2	5	96	128875	1288.751	11.1%
16	2	8	60	207265	809.627	8.8%
20	2	10	48	348623	871.558	9.2%
32	2	16	30	537840	525.2344	7.1%
40	2	20	24	863245	539.5283	7.2%
80	2	40	12	2417935	377.8023	6.0%
3	3	1	320	22698.59	2522.066	15.6%
6	3	2	160	68350.29	1898.619	13.5%
12	3	4	80	145750.9	1012.159	9.9%
15	3	5	64	234999.9	1044.444	10.0%
24	3	8	40	386651.4	671.2697	8.0%
30	3	10	32	691436.6	768.2628	8.6%
48	3	16	20	998596.5	433.4186	6.5%
60	3	20	16	1717301	477.028	6.8%
120	3	40	8	4645044	322.5725	5.6%
4	4	1	240	33835.5	2114.719	14.3%
8	4	2	120	102048.1	1594.502	12.4%
16	4	4	60	206941.1	808.3637	8.8%
20	4	5	48	335279.7	838.1992	9.0%
32	4	8	30	482145.1	470.8448	6.7%
40	4	10	24	884015.9	552.51	7.3%
64	4	16	15	1092044	266.6123	5.1%
120	6	20	12	1829737	127.0651	3.5%
240	6	40	4	4388502	76.18926	2.7%
6	6	1	160	61521.27	1708.924	12.8%
12	6	2	80	186091	1292.299	11.2%
24	6	4	40	338031.5	6771.037	7.5%
30	6	5	32	590421.1	656.0234	7.9%
48	6	8	20	813136.5	352.9238	5.8%
60	6	10	16	1579341	438.7059	6.5%

Table 4 ctd.

Plot size and shape

Size, m^2	Width, m	Length, m	Total no of plots	V(x)	Vx	CV
96	6	16	10	1629653	176.8286	4.1%
120	6	20	8	3286099	228.2013	4.7%
240	6	40	4	4293629	74.54217	2.7%
12	12	1	80	169661.8	1178.207	10.6%
24	12	2	40	521244.7	904.9387	9.3%
48	12	4	20	793000.3	344.1842	5.8%
60	12	5	16	1724759	479.0998	6.8%
96	12	8	10	2003785	217.4246	4.6%
120	12	10	8	4620347	320.8574	5.6%
192	12	16	5	4046890	109.7789	3.3%
240	12	20	4	9923809	172.2883	4.1%
480	12	40	2	18343625	79.61643	2.8%

Table 5 Tests for homogeneity of variance for plots of the same size but different shapes

Grain yield broadcast sown

Size, m^2	W, m	L, m	No of plots	$V_{(x)}$	$Log(V_{(x)})$	F	Chi Square
2	1	2	360	19004	4.278856	1.105806	
	2	1	360	17186	4.235177		
3	1	3	240	36835	4.566255	1.123542	
	3	1	240	32784.31	4.515666		
4	1	4	180	60376	4.780862		0.463496
	2	2	180	58565	4.767638		
	4	1	180	54644.97	4.73755		
6	1	6	120	119074	5.075818		0.447111
	2	3	120	116618	5.066766		
	3	2	120	113329.7	5.054344		
	6	1	120	106068	5.025584		
8	2	4	90	195104.1	5.290266	1.025224	
	4	2	90	190303.9	5.279448		
10	1	10	72	289482	5.461621	1.034847	
	2	5	72	279734	5.446745		
12	1	12	60	412081	5.614983		0.187301
	2	6	60	391046	5.592228		
	3	4	60	384533	5.584934		
	4	3	60	386725.1	5.587402		
	6	2	60	379131.4	5.57879		
	12	1	60	370808.6	5.56915		
15	1	15	48	614519	5.788535	1.111533	
	3	5	48	552857.1	5.742613		
20	1	20	36	1051833	6.021947		0.092091
	2	10	36	984702.4	5.993305		
	4	5	36	950317.6	5.977869		
24	2	12	30	1379444	6.139704		0.007047
	4	6	30	1362210	6.134244		
	6	4	30	1345487	6.128879		
	12	2	30	1383762	6.141061		
30	1	30	24	2003641	6.30182		0.04403
	2	15	24	2118932	6.326117		
	3	10	24	1975598	6.295699		
	6	5	24	1954844	6.291112		
36	3	12	20	2734216	6.436833		2.317429
	6	6	20	2822852	6.450688		
	12	3	20	5039121	6.702355		
40	2	20	18	3748983	6.573914	1.05689	
	4	10	18	3547183	6.549884		
48	4	12	15	4971045	6.696448	1.866014	
	12	4	14	2663991	6.425533		
60	2	30	12	7189975	6.856727		0.030242
	4	15	12	7927635	6.899144		
	6	10	12	7310618	6.863954		
	12	5	12	7598651	6.880736		
72	6	12	10	10362290	7.015456	0.917626	
	12	6	10	11292493	7.05279		
90	3	30	8	14584639	7.163896	0.880924	

Table 5 ctd.

Size, m^2	W, m	L, m	No of plots	$V_{(x)}$	$Log(V_{(x)})$	F	Chi Square
	6	15	8	16556065	7.218957		
120	4	30	6	27948644	7.446361		0.017546
	6	20	6	30769216	7.488116		
	12	10	6	31395922	7.496873		
180	6	30	4	60747614	7.783529	0.800912	
	12	15	4	75848055	7.879944		

Table 6 Tests for homogeneity of variance for plots of the same size but different shapes
Grain yield row planted

Size, m^2	W, m	L, m	No of plots	$V_{(x)}$	$Log(V_{(x)})$	F	Chi Square
2	1	2	480	13124	4.118066	1.070852	
	2	1	480	12256	4.088337		
4	1	4	240	32356	4.509952		0.517731
	2	2	240	35514	4.550397		
	4	1	240	33835.5	4.529373		
6	3	2	160	68350.29	4.83474	1.111003	
	6	1	160	61521.27	4.789025		
8	1	8	120	89363	4.951157		1.727541
	2	4	120	80222	4.904293		
	4	2	120	102048.1	5.008805		
10	1	10	96	140710	5.148324	1.091831	
	2	5	96	128875	5.110169		
12	3	4	80	145750.9	5.163611		1.182644
	6	2	80	186091	5.269725		
	12	1	80	169661.8	5.229584		
16	1	16	60	237487	5.37564		0.371227
	2	8	60	207265	5.316525		
	4	4	60	206941.1	5.315847		
20	1	20	48	362132	5.558867		0.069074
	2	10	48	348623	5.542356		
	4	5	48	335279.7	5.525407		
24	3	8	40	386651.4	5.58732		1.93528
	6	4	40	338031.5	5.528957		
	12	2	40	521244.7	5.717042		
30	3	10	32	691436.6	5.839752	1.171091	
	6	5	32	590421.1	5.771162		
32	2	16	30	537840	5.730653	1.115515	
	4	8	30	482145.1	5.683178		
40	1	40	24	963494	5.983849		0.076055
	2	20	24	863245	5.936134		
	4	10	24	884015.9	5.94646		
48	3	16	20	998596.5	5.99939	1.259264	
	12	4	20	793000.3	5.899273		

Table 6 ctd.

Size, M^2	W, m	L, m	No of plots	$V_{(x)}$	$Log(V_{(x)})$	F	Chi Square
60	3	20	16	1717301	6.234846		0.035564
	6	10	16	1579341	6.198476		
	12	5	16	1724759	6.236729		
96	6	16	10	1629653	6.212095	0.813287	
	12	8	10	2003785	6.301851		
120	3	40	8	4645044	6.66699		1.707016
	6	20	12	1829737	6.262389		
	6	20	8	3286099	6.516681		
	12	10	8	4620347	6.664675		
240	6	40	6	4388502	6.642316		1.118343
	12	20	4	9923809	6.996678		

Using this and the previous results, the fitted models are (see equations (1), (4), and (5)) .

Row planted $$V_x = \frac{V_1}{x^{.62378}}$$

Broadcast Sown $$V_x = \frac{V_1}{x^{.2170}}$$

The relatively small value of the estimate of the Smith's index of soil heterogeneity, b for broadcast sown plots, 0.2170, indicates that there is a high degree of correlation between adjacent plots and hence, the change in fertility gradient over the experimental farm is gradual than in patches. For row planted, however, the estimated b value, 0.62378, shows that there is more than average soil heterogeneity over the experimental farm through this method of plantation. That is to say, fertile plots occur relatively dispersed over the farm than the uni-directional fertility gradient

observed in broadcast sown method. Both results are strongly justified by the productivity contour maps of Figures.1 and 2.

The reasons for the observed difference in the degree of adjacent plot correlation as measured by the Smith's, b, for the two crop geometry's (plantation methods) are many. Girma and Amsale (1999) discussed three causes as 1) severe competition between neighboring plants resulted in high degree of adjacent plant correlation and hence lower value of the heterogeneity coefficient. 2) Plant density per plot cannot be controlled and is different from plot to plot which exaggerates the effect of plant geometry. 3) The amount of fertilizer applied to different plots is not the same and this gives rise to different rate of fertilizer competition. It is plausible to believe the reasons from experience. However, justifying the relatively high observed difference among the estimated b values as entirely due to the difference from crop geometry can't be expected in this study. Because apart from the difference in crop geometry ., the changes could be due to the native soil fertility variation (change in the nutrient concentration) over the experimental sites taken or may be the cultural and management practices were not kept constant. The Difference in the amount of fertilizer applied, given as the third reason, does not seem to be reasonable as fertilizer application should be independent of plantation method used.

3.1.4 Plot shape

To assess the pattern of reduction in variance between yields from different plots due to plot shape, the method by Narayana and Ramanata (1983) is adopted. All plots will be classified into three groups accordingly to each of the restriction, $x_1<x_2$, $x_1=x_2$, $x_1>x_2$ where x_1 and x_2 are the length and breadth of the plots in terms of the number of basic units residing along and across the crop rows. The heterogeneity coefficient, b, will be computed and compared for each group using the usual way of Smith's empirical variance law.

The estimated b values for the three regions are 0.619532, 0.60337, and 0.58298 for row planted and 0.230769, 0.24938, 0.277352 for broadcast sown respectively. In both plantation methods, there is no significant difference of the computed values for the three regions to favor long narrow rectangular plots. But, the relatively large value of the b coefficient for $x_1<x_2$ in row planted and $x_1>x_2$ for broadcast sown method shows that rectangular plots with smaller number of basic units along the row than across the column are more efficient for row planted and vice versa than the other shapes. This results were also confirmed by Girma and Amsale (1999) using the improved methods of Narayana and Ramanata (1985).

3.1.5 Plot Size

As discussed in the introductory part of this paper, the experimental error should be kept to the minimum to measure the true difference between treatments and their levels under study. As had been pointed out earlier, a reduction in this experimental error can be achieved either by increasing the number of replication for a fixed plot size or increasing the plot size for a fixed number of replications. The size of a plot is said to be optimum when the experimental error is minimum for a fixed cost and replication.

Whenever cost estimates are available, optimum plot size can be found using the equation (Mohamed 1991)

$$x_{opt} = bk_1/(1-b)k_2$$

where b= Coefficient of soil heterogeneity
k_1 = Cost per plot
k_2 = Cost per unit area of test area.

In the study of uniformity trials, when cost estimates are not available, the method widely applied to find optimum plot size is the one by Hathway (1961), a method which is also adopted by Girma and Amsale for the same data. In this study we use a relatively easy method, known as 'comparable variance method', due to Keller as cited in Narayana and Ramanta (1983).

The variance of basic units will be divided by the number of basic units per plot to give comparable variance, V. In order to arrive at optimum plot size, variances between basic units are assumed to give complete or 100% information. The Relative Information, $RI = \frac{V}{V_1} X\ 100$ for each of the number of combinations of different plot sizes and shapes will be calculated. If the occurrence of plots is completely at random, (complete heterogeneity), then b = 1 and from the empirical variance low of Smith (1) we have,

$$xV_x = V_1 = V \quad .$$

Hence, under the assumption of complete randomness, the comparable variance is constant and any plot will provide a complete relative information. This situation seldom exists because of adjacent plot correlation. So, as the plot size increases, both the variance V_x, and relative information, RI, decrease. The change in RI and the average relative information over all possible plot sizes and shapes will be used to derive an optimum plot size.

Tables 7 and 8 show the Variance per unit area, V_x, Comparable Variance, V, and RI, for broadcast and row plantation methods respectively.

Table 7 Variance per unit area, V_x, comparable variance, V, and relative Information, RI, for plots of different size using Broadcast sowing method

Plot Size, m^2	V_x	V	RI
1	5884	5884	100
2	4524	9048	65
3	3868	11603	51
4	3616	14465	41
5	3490	17449	34
6	3160	18962	31
8	3011	24088	24
9	2783	25051	23
10	2846	28461	21
12	2690	32282	18
15	2594	38913	15
16	2625	42001	14
18	3216	57881	10
20	2489	49781	12
24	2375	56989	10
30	2237	67108	9
36	2725	98113	6
40	2280	91202	6
45	8102	364599	2
48	1657	79532	7
60	2094	125611	5
72	2089	150380	4
80	2231	178474	3
90	1922	173004	3
120	2086	250316	2
144	2213	318726	2
180	2108	379432	2
240	2623	629595	1
360	2776	999343	1

Table 8 Variance per unit area, V_x, comparable variance, V, and relative information RI, for plots of different size using Row plantation method

Plot Size, m^2	Vx	V	RI
1	4563.226	4563	100
2	3172.46	6345	72
3	2522.066	7566	60
4	2118.856	8475	54
5	2014.159	10071	45
6	1803.772	10823	42
8	1414.754	11318	40
10	1347.924	13479	34
12	1160.888	13931	33
15	1044.444	15667	29
16	848.5585	13577	34
20	871.6955	17434	26
24	629.0649	15098	30
30	712.1431	21364	21
32	498.0396	15937	29
40	564.7407	22590	20
48	376.8422	18088	25
60	464.9446	27897	16
64	266.6123	17063	27
80	331.8494	26548	17
96	197.1266	18924	24
120	290.5437	34865	13
160	171.4258	27428	17
192	109.7789	21078	22
240	123.4153	29620	15
480	79.61643	38216	12

There is more loss in relative information in broadcast sown fields than in row planted ones because of greater adjacent plot correlation in the former. The average relative information that can be gained over all possible combinations of basic units is 18% which corresponds with plot size of 12 m^2 for broadcast sown fields. Plots of the same size (12 m^2) gave a relative information of 34% using row plantation method with an average relative information of 33%. Therefore, a plot size of 12 m^2 for broadcast

sown and 8 to 12 m^2 for row planted with rectangular shapes as discussed earlier seems to be optimum.

3.2 Spatial autocorrelation

Smith's index of soil heterogeneity, b, measures the degree of correlation between adjacent plots based on the empirical relation of variance per basic unit area and size of plots. It is not however clear how plots have to be close to each other to be termed as 'adjacent'. The Moran's, I, and Geary's, c, coefficients are relatively easy and direct measures of spatial autocorrelation between 'neighboring' plots in an experimental site. Besides the fact that the coefficients take the usual form of autocorrelation, their distribution is well studied so that they can be used for testing significance of spatial autocorrelation in neghbouring plots or counties in a study area. The term 'neighbor' can be defined in many ways and different weights can be associated with different counties, making the coefficients to be applied in a surprisingly wide area of disciplines.

Under the null hypothesis stating that there exists no spatial autocorrelation between yields of neighboring plots, Moran's, I, and Geary's, c, coefficients were computed for plots of different sizes and shapes (totals have been taken in simulating plots of different sizes and shapes). Since the effect of plot shape was not found to be significant by the homogeneity of variance tests under both plantation methods, the statistics are computed for different size but for randomly selected shape. As, the connection matrix can take any of the three forms discussed, namely Rook's case,

Bishop's case, and Queen's case, to represent neighborhood within plots, results of the uniformity trial data under both plantation for all the cases will be presented in the subsequent Tables 8-13.

To test for significance, the moments of the coefficients were computed under both the Normality and Randomization assumptions as given in the methodological part of this study. The sample kurtosis coefficient, b_2, for most of the sizes considered is nearly equal to 3, the normal curve value, under all the three cases. From Tables 8 to 10, we see that the standard deviation of the coefficients are not significantly different under the two assumptions, implying that the normality assumption does not seriously affect the results.

Results for the Rook's case (Tables 8 and 9) show that there is a significant positive spatial autocorrelation between neighboring plots using broadcast sowing method for all sizes considered. This shows that there is an overall high adjacent plot correlation throughout the experimental farm, arising either from the naive soil fertility gradient or from the plantation method used. For row planted fields however, both coefficients show significant positive spatial autocorrelation between yields of plots up to a simulated plot of 24 basic units (24 m^2). Hence, for row planted fields, taking plots of size 12 m^2 or more will minimize and ultimately eliminate the observed spatial autocorrelation between adjacent plots after a plot size of about 24 m^2,.

Tables 10 and 11 show results for Bishop's case, where plots are said to be 'neighbors' if they are connected with a vertex. As with the Rook's case, for

broadcast sown fields all the coefficients were significant tested as standard normal deviates implying that there is a significant spatial autocorrelation between adjacent plots connected at a vertex. For row planted fields, the Moran's I coefficient shows a significant spatial autocorrelation for plots of size 4 m^2. Both the Moran's and Geary's coefficients show a significant negative spatial autocorrelation between plots of size 30 m^2 (3 by 10). We can see from this that there is no a diagonal fertility trend in all directions for row planted fields.

Table 8 Results for I and c using broadcast sowing method
Rook's case

Size, m^2	W, m	L, m	Total No. Plots		Moran's I		Geary's c	
					N	R	N	R
4	2	2	240	Expected Value under Ho	-0.005	-0.00	1	1
				Observed Value	0.614	0.61	0.36582	0.365826
				Standard deviation	0.054	0.055	0.05685	0.055189
				Standard normal deviate	11.28	11.222	11.1549	11.49097
				Sig.	**	**	**	**
8	2	4	80	Expected Value under Ho	-0.011	-0.011	1.0	1.0
				Observed Value	0.6554	0.655	0.30029	0.300299
				Standard deviation	0.0776	0.078	0.08051	0.079199
				Standard normal deviate	8.5874	8.541	8.68212	8.834701
				Sig.	**	**	**	**
16	4	4	40	Expected Value under Ho	-0.022	-0.025	1.0	1.0
				Observed Value	0.730	0.730	0.27933	0.279334
				Standard deviation	0.1132	0.114	0.11717	0.11532
				Standard normal deviate	6.6485	6.577	6.15231	6.249257
				Sig.	**	**	**	**
60	6	10	32	Expected Value under Ho	-0.090	-0.090	1.0	1.0
				Observed Value	0.7558	0.755	0.12845	0.128459
				Standard deviation	0.2163	0.230	0.21088	0.215048
				Standard normal deviate	3.9140	3.679	4.13284	4.05277
				Sig.	**	**	**	**
120	12	10	24	Expected Value under Ho	-0.2	-0.2	1.0	1.0
				Observed Value	0.645165	0.645	0.150903	0.150903
				Standard deviation	0.357771	0.400	0.338062	0.352175
				Standard normal deviate	2.362308	2.108	2.511663	2.41101
				Sig.	**	**	**	**

** Significant at 5% level

Table 9 Results for I and c using row plantation method

Rook's Case

Size, m^2	W, m	L, m	Total no. Plots		Moran's I coefficient		Geary's c coefficient	
					N	R	N	R
4	2	2	240	Expected Value under Ho	-0.00418	-0.00418	1.0	1.0
				Observed Value	0.2537	0.2537	0.69953	0.699533
				Standard deviation	0.047595	0.047441	0.04916	0.05034
				Standard deviate	5.419061	5.436654	6.11138	5.968756
				Sig.	**	**	**	**
12	3	4	80	Expected Value under Ho	-0.01266	-0.01266	1.0	1.0
				Observed Value	0.251727	0.251727	0.69737	0.697379
				Standard deviation	0.083781	0.083818	0.08706 3	0.086952
				Standard deviate	3.155684	3.154294	3.47590	3.480337
				Sig.	**	**	**	**
24	3	8	40	Expected Value under Ho	-0.02564	-0.02564	1.0	1.0
				Observed Value	0.242979	0.242979	0.56866	0.568665
				Standard deviation	0.117405	0.117321	0.12240	0.122541
				Standard deviate	2.287972	2.28961	3.52376	3.519937
				Sig.	**	**	**	**
30	3	10	32	Expected Value under Ho	-0.03226	-0.03226	1.0	1.0
				Observed Value	0.127007	0.127007	0.87091	0.870913
				Standard deviation	0.13065	0.132092	0.13639	0.134505
				Standard deviate	1.219021	1.205715	0.94643	0.959721
				Sig.				
40	4	10	24	Expected Value under Ho	-0.04348	-0.04348	1.0	1.0
				Observed Value	0.110782	0.110782	0.83502	0.835022
				Standard deviation	0.152134	0.149872	0.15759	0.159318
				Standard deviate	1.013978	1.029279	1.04685	1.035525
				Sig.				

** Significant at 5% level

Table 10 Results for I and c using broadcast sowing method
Bishop's case

Size, m^2	W, m	L, m	Total no. Plots		Moran's I		Geary's c	
					N	R	N	R
4	2	2	180	Expected Value under Ho	-0.005586	-0.0055	1.0	1.0
				Observed Value	0.43755	0.437	0.51851	0.518518
				Standard deviation	0.05808796	0.058	0.06652	0.059112
				Standard normal deviate	7.62866797	7.59112	7.23774	8.145194
				Sig.	**	**	**	**
8	2	4	90	Expected Value under Ho	-0.01123596	-0.0112	1.0	1.0
				Observed Value	0.43628685	0.4362	0.47217	0.472173
				Standard deviation	0.08268507	0.0831	0.09622	0.090034
				Standard normal deviate	5.41237775	5.383	5.48547	5.862501
				Sig.	**	**	**	**
16	4	4	45	Expected Value under Ho	-0.02273	-0.022	1.0	1.0
				Observed Value	0.61419	0.614	0.41231	0.412314
				Standard deviation	0.12859	0.129	0.1517	0.141282
				Standard normal deviate	4.95320	4.900	3.87385	4.159653
				Sig.	**	**	**	**
30	6	10	12	Expected Value under Ho	-0.0909091	-0.0909	1.0	1.0
				Observed Value	0.6212642	0.6212	0.19465	0.194651
				Standard deviation	0.2884512	0.3066	0.28823	0.288946
				Standard normal deviate	2.4689562	2.3226	2.79411	2.787195
				Sig.	**	**	**	**

** Significant at 5% level

Table 11 Results for I and c using row plantation method
Bishop's case

Size, m^2	W, m	L, m	Total No. Plots		Moran's I c		Geary's c	
					N	R	N	R
4	2	2	240	Expected Value under Ho	-0.00418	-0.004184	1.0	1.0
				Observed Value	-0.1385	-0.138501	0.932778	0.9327
				Standard deviation	0.050163	0.0500012	0.057102	0.0503
				Standard normal deviate	-2.67761	-2.686283	1.177234	1.3353
				Sig.	**	**		
12	3	4	80	Expected Value under Ho	-0.01266	-0.012658	1.0	1.0
				Observed Value	-0.10156	-0.101562	0.826853	0.8268
				Standard deviation	0.091488	0.0915278	0.107849	0.1073
				Standard normal deviate	-0.97175	-0.971328	1.605462	1.6132
				Sig.				
24	3	8	40	Expected Value under Ho	-0.02564	-0.025641	1.0	1.0
				Observed Value	-0.23639	-0.236395	1.187515	1.1875
				Standard deviation	0.129738	0.1296473	0.155852	0.1565
				Standard normal deviate	-1.62445	-1.625592	-1.20316	-1.1981
				Sig.				
30	3	10	32	Expected Value under Ho	-0.03226	-0.032258	1.0	1.0
				Observed Value	-0.35221	-0.352208	1.348631	1.34863
				Standard deviation	0.145307	0.1468704	0.17604	0.16633
				Standard normal deviate	-2.20189	-2.178453	-1.98040	-2.09591
				Sig.	**	**	**	**
40	4	10	24	Expected Value under Ho	-0.04348	-0.043478	1.0	1.0
				Observed Value	-0.24803	-0.248028	1.111422	1.11142
				Standard deviation	0.175494	0.1729746	0.210684	0.22104
				Standard normal deviate	-1.16557	-1.182544	-0.528859	-0.5040
				Sig.				

** Significant at 5% level

Table 12 Results for I and c Using, broadcast sowing method
Queen's case

Size, m^2	W, m	L, m	Total no. Plots		Moran's I		Geary's c	
					N	R	N	R
4	2	2	180	Expected Value under Ho	-0.005587	0.00558	1.0	1.0
				Observed Value	0.53099	0.53099	0.437944	0.437944
				Standard deviation	0.039501	0.03969	0.045918	0.040286
				Standard normal deviate	13.58366	13.5168	12.24049	13.95169
				Sig.	**	**	**	**
8	2	4	90	Expected Value under Ho	-0.011236	-0.01123	1.0	1.0
				Observed Value	0.552828	0.55282	0.380775	0.380775
				Standard deviation	0.055365	0.05566	0.065901	0.061114
				Standard normal deviate	10.18813	10.1340	9.396281	10.13235
				Sig.	**	**	**	**
16	4	4	45	Expected Value under Ho	-0.022727	-0.02272	1.0	1.0
				Observed Value	0.679367	0.67936	0.337513	0.337513
				Standard deviation	0.081574	0.08244	0.098821	0.091099
				Standard normal deviate	8.60687	8.51622	6.703900	7.272178
				Sig.	**	**	**	**
30	6	10	12	Expected Value under Ho	-0.090909	-0.0909	1.0	1.0
				Observed Value	0.7	0.7	0.153917	0.153917
				Standard deviation	0.148825	0.15783	0.16248	0.153588
				Standard normal deviate	5.3418889	5.03696	5.207315	5.508775
				Sig.	**	**	**	**
40	12	10	6	Expected Value under Ho	-0.2	-0.2	1.0	1.0
				Observed Value	0.645165	0.64516 5	0.150903	0.150903
				Standard deviation	0.357771	0.40093	0.338062	0.352175
				Standard normal deviate	2.362308	2.10810	2.511663	2.411010
				Sig.	**	**	**	**

** Significant at 5% level

Table 13 Results for I and c, using row plantation method
Queen's case

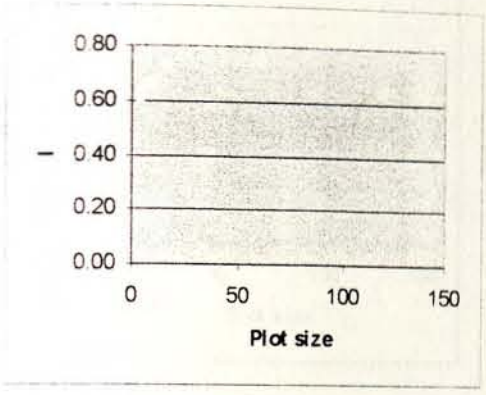
Size, m^2	W, m	L, m	Total No. Plots		Moran's I		Geary's c	
					N	R	N	R
4	2	2	240	Expected Value under Ho	-0.00418	-0.00418	1.0	1.0
				Observed Value	0.06775	0.06775	0.859944	0.859948
				Standard deviation	0.034273	0.03416	0.039569	0.043204
				Standard normal deviate	2.098852	2.10565 1	3.539495	3.241694
				Sig.	**	**	**	**
12	3	4	80	Expected Value under Ho	-0.01266	-0.01266	1.0	1.0
				Observed Value	0.09063	0.09063	0.829182	0.829182
				Standard deviation	0.06034	0.06037	0.072596	0.072207
				Standard normal deviate	1.71168	1.71093 4	2.353003	2.365634
				Sig.			**	**
24	3	8	40	Expected Value under Ho	-0.02564	-0.02564	1.0	1.0
				Observed Value	0.02726	0.02726	0.847148	0.847145
				Standard deviation	0.082748	0.08269	0.103299	0.103804
				Standard normal deviate	0.63931	0.63975 7	1.47971	1.47250
				Sig.				
30	3	10	32	Expected Value under Ho	-0.03226	-0.03226	1.0	1.0
				Observed Value	-0.08711	-0.08711	1.084361	1.084361
				Standard deviation	0.09099	0.09195	0.11571	0.108023
				Standard normal deviate	-0.60287	-0.59652	-0.72907	-0.78095
				Sig.				
40	4	10	24	Expected Value under Ho	-0.04348	0.04348	1.0	1.0
				Observed Value	-0.04378	-0.04378	0.954087	0.95408
				Standard deviation	0.105506	0.10417	0.13412	0.14225
				Standard normal deviate	-0.00288	-0.00292	0.342330	0.322661
				Sig.				

** Significant at 5% level

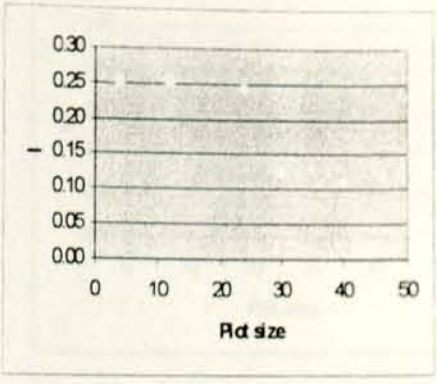
For Queen's case, Tables 12 and 13 show intermediate results of the above two cases. The same result holds for broadcast sown fields, significant positive spatial autocorrelation between adjacent plots in any way within the experimental farm through this plantation method. The Geary's c coefficient shows significant positive spatial autocorrelation between neighboring plots up to a plot size of 12 m^2 (3 by 4), while using the Moran's I statistic, the observed spatial autocorrelation was significant only between simulated plots of 4 basic units (4 m^2).

From the above discussion, the significant spatial autocorrelation as measured by the I or c statistics for all plot sizes considered using broadcast sowing method in all cases, clearly shows that there is a high degree of spatial autocorrlation between yield of adjacent plots and there exists a continuos fertility gradient over the experimental farm. This result was also confirmed using the results of uniformity trials as a very small value of Smith's index of soil heterogeneity and the uni-directional fertility gradient from the productivity contour map of Figure 1. Especially from the contour map, we see that plots residing to the left side of the experimental site with size approximately of 240 m^2 (12 by 20), produce very less, and the products of plots of approximately of the same size in the central part are more similar than the right extremes which are highly fertile. Hence, even increasing the size of the plots will not generally make the observed spatial autocorrelation insignificant.

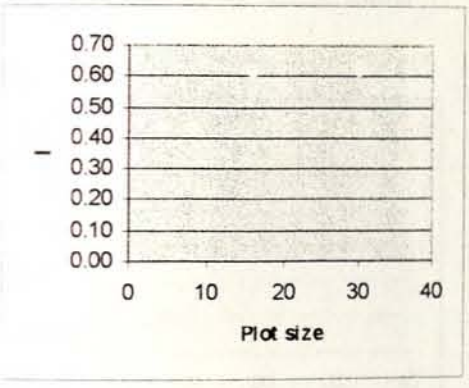
Plot 1 Bisop's case, broadcast sown



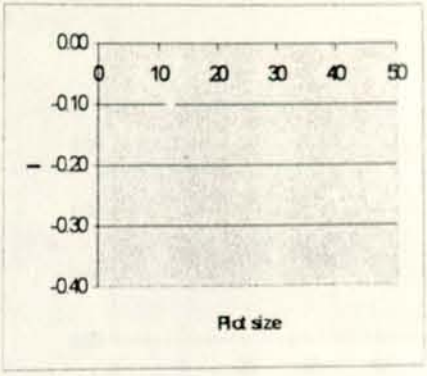
Plot 2 Bisop's case, row planted



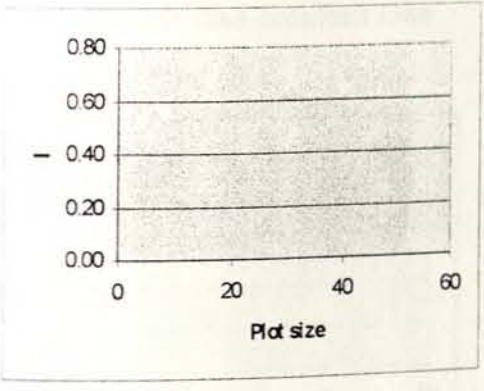
Plot 3 Rook's case, broadcast sown



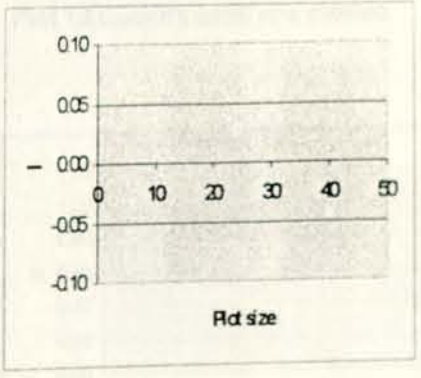
Plot 4 Rook's case, row planted



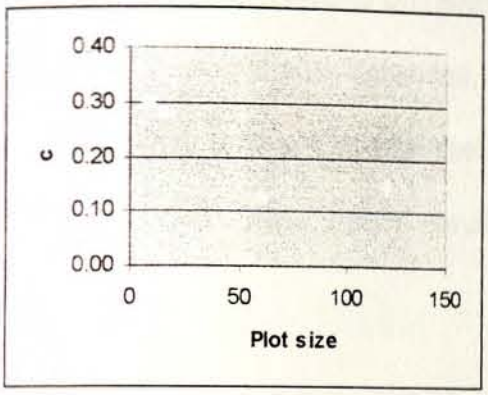
Plot 5 Queen's case, broadcast sown



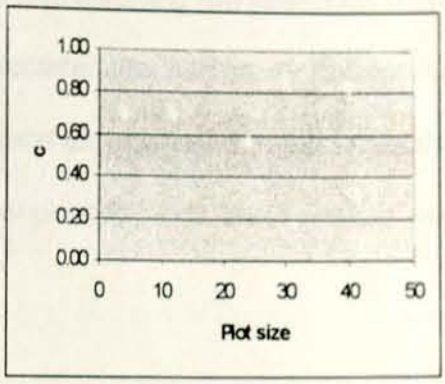
Plot 6 Queen's case, planted



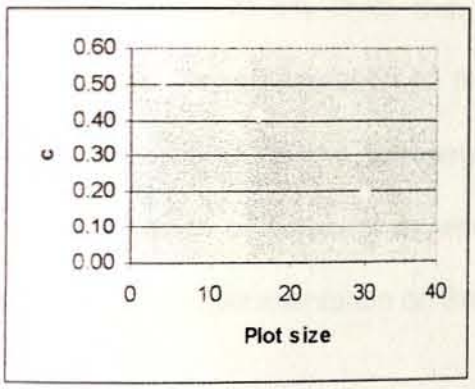
Plot 7 Bisop's case, broadcast sown



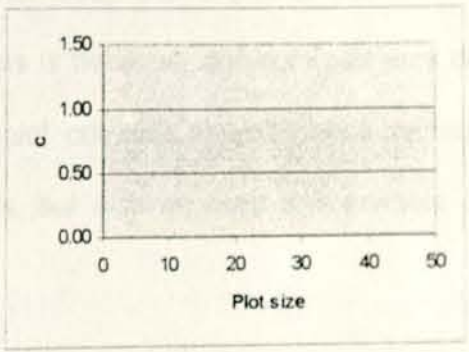
Plot 8 Bisop's case, row planted



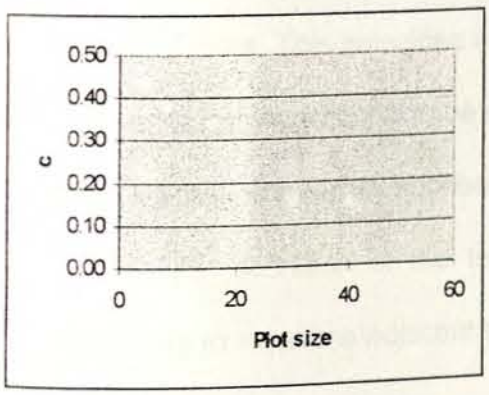
Plot 9 Rook's case, broadcast sownd



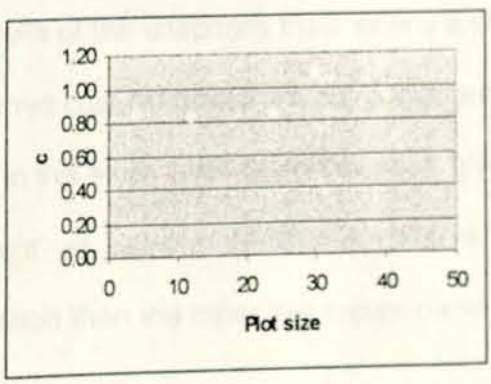
Plot 10 Rook's case, row planted



Plot 11 Queen's case, broadcast sown



Plot 12 Queen's case, row planted



Plots 1-12 show values of Moran's I and Geary's c coefficients calculated under the three different assumptions on neighborhood. From involving row plantation methods, we see that as plot size increases, spatial autocorrelation decreases (except for plot of size 30 m^2), as might be expected from practical experience. But, this is not true for broadcast sown fields, hence results from row planted fields show a more realistic situation.

Using row plantation method, Rook's case, for a plot of size 30 m^2 (3 x 10), both the tests confirmed that there is no significant spatial autocorreltion between yields of the simulated plots at 5% level. But, this does not mean that a plot of size 30 m^2 is optimum for experimentation on that area. This is because, optimum plot size does not only depend on the between adjacent plot correlation, and hence variability between yields of adjacent experimental plots, but also on fixed and variable cost estimates of experimentation on that site.

For Queen's case, a plot size of about 4 m^2 (using Moran's, I) and 12 m^2 (using Geary's, c) minimizes the observed spatial autocorrelation between the yields of 'neighboring' plots. This coincides with the results of the uniformity trials where a plot size of 8 to 12 m^2 were found to be optimum when cost estimates are not considered. This implies that, the term "neighbor' as used in this study (with Queen's case), gives a more similar definition of the term 'adjacent' as used in Smith's coefficient of heterogeneity to measure adjacent plot correlation than the other two cases namely: Bishop's case, and Rook's case.

3.3 Conclusion and recommendation

The results of this study demonstrate that within the same experimental farm, soil fertility pattern may vary with the type of plantation method used. This is due to the spatial correlation between yield of adjacent experimental plots or due to the naïve spatial patterns of fertility parameters of the soil over the experimental farm. Although this study characterized the soil fertility pattern at a single experimental farm of the IAR research site- Holleta, the results illustrate some useful concepts in characterizing soil fertility patterns:

i) Constructing a productivity contour map can be advantageous to have a general visual inspection of the experimental farm and especially to choose appropriate block shapes, if the plot size is determined.

ii) If not for time and money, studying spatial patterns of soil fertility parameters like nutrient concentration using the methods of uniformity trials or spatial statistical methods will be an advantage besides studying the fertility pattern to choose site specific crop management.

iii) Using spatial autocorrelation measures with the widely used uniformity trial techniques, enables besides testing for significance of the observed spatial correlation between adjacent plots to study the pattern of the observed autocorrelation within plots of different sizes and shapes.

Finally, the following problems can further be studied on the same data:

- Pattern of spatial autocorrelation can be studied between yield of plots having the same size but different shapes.

- Two dimensional spectral or correlelogram analysis can be done to see how the autocorrelation diminishes in space.

- Different spatial statistical models can be fitted to the data and compared.

APPENDIX 1

MODELS FOR SPATIAL DATA

a. Why spatial models?

The thinking that data that are close together, in time or space, are likely to be correlated (i.e., cannot be modeled as statistically independent) is a natural one and has been used by statisticians in modeling physical or social phenomenon for years.

In the early 1920s and 1930s, R.A. Fisher was clearly aware of spatial dependencies in agricultural field experiments. He established the principles of blocking, randomization, and replication to control unwanted bias and to neutralize the effect of spatial autocorrelation.

Smith (1938), recognizing the presence of spatial correlation in field experiments, was concerned with choosing plot dimensions so that any increase in plot size would result in little decrease in error variance.

Nearest neighbor methods for analyzing agricultural field trials attempt to take spatial dependencies into account, indirectly by using residuals from neighboring plots as covariates or by differencing (Papadakins (1937),Bartlett (1967), Besag (1974), Besag and Kempton (1986)).

Generally, when dealing with data where space and/or time interactions are likely, two approaches can be used. Departures from independence paradigm could be modeled, or statistical procedures can be constructed that would be robust to these departures. The former relates to spatial models, and will be discussed here.

b. The general spatial model

A general model for spatial data, which has a very simple structure and is flexible enough to handle an extremely large class of problems as given in Cressie (1991) will be presented.

Let $\mathbf{s} \in IR^d$ be a generic data location in d- dimensional euclidean space. Suppose the potential datum $\mathbf{Z}(\mathbf{s})$ at spatial location $\mathbf{s}$ is a random quantity. Assume that $\mathbf{s}$ vary over the index set $D \subset IR^d$, so as to generate the multivariate random process:

$$\{ \mathbf{Z}(\mathbf{s}) : \mathbf{s} \in D \} \qquad (5)$$

where a realization of (5) is denoted by $\{ \mathbf{z}(\mathbf{s}) : \mathbf{s} \in D \}$ (6)

At the very least, the above general spatial model is used to summarize extant data or to predict unobserved data of any spatial nature.

Usually, D is assumed fixed (i.e. not random) subset of IR^d, but it can more generally be assumed to vary from realization to realization. The simple structure of (5) and this more general assumption on D and **Z**, helps to represent all the three different kinds of spatial data discussed in Chapter 1. Specifically:

Point patterns: D is a point process in IR^d or subset of IR^d; $\mathbf{z}(\mathbf{s})$ is a random vector at location $\mathbf{s} \in D$. In its most general form, this generates a marked spatial point process; when no **z** is specified, the usual spatial point process is obtained (or one could think of scaling $\mathbf{z}(\mathbf{s})=1$, for all $\mathbf{s} \in D$.)

Lattice data: D is fixed (regular or irregular) collection of countabley many points of IR^d; $\mathbf{z}(\mathbf{s})$ is a random vector at location $\mathbf{s} \in D$.

Geostatistical data: D is fixed subset of IR^d that contains a d- dimensional rectangle of positive volume; $\mathbf{z}(\mathbf{s})$ is a random vector at location $\mathbf{s} \in D$

To modify the above general spatial model (5), so that we can represent spatial temporal data, let the (multivariate) time series process at time t be represented by:

$$\{ \mathbf{Z}(t) : -\infty < t < \infty \} \qquad (7)$$

A space –time process can be denoted as:

$$\{\mathbf{Z}(\mathbf{s},t) : \mathbf{s} \in D, t \in T\} \qquad (8)$$

where each of Z, D , and T are possibly random.

c. Nearest-neighbor system on a lattice

There are two main approaches to the specification of spatial stochastic process on a lattice. This steams from the non equivalent definition a 'nearest –neighbor' system originally due to Whittle (1963) and Barttlet (1955) as cited in Besag (1974). For definiteness, consider a uniformly spaced rectangular lattice with sites labeled (i,j) and the associated random variable $\{ X_{i,j} \}$.

According to Whittle, the joint probability distribution of the random variables should be of the form

$$\prod_{i,j} Q_{i,j}(x_{i,j}; x_{i-1,j}, x_{i-1,j}, x_{i,j-1}, x_{i,j-1}) \qquad (9)$$

where $x_{i,j}$ is the observed value of the random variable, $X_{i,j}$, and $Q_{i,j}$ its density function. On the other hand, Bartlett's definition requires the conditional probability distribution of $x_{i,j}$, given all other site values, to depend only upon the four nearest sites to (i,j), namely $x_{i-1,j}$, $x_{i+1,j}$, $x_{i,j-1}$, and $x_{i,j+1}$.

Bartlett (1967) and Besag (1974) discuss the advantages and disadvantages of different spatial models based on the above basic definitions of 'nearest–neighbor' systems with examples.

Based on the Whittle's definition, a first order simultaneous autoregressive model can be defined as :

$$X_{i,j} = \beta_1 X_{i-1,j} + \beta_1' X_{i+1,j} + \beta_2 X_{i,j-1} + \beta_2' X_{i,j+1} + U_{i,j} \quad . \qquad (10)$$

where { Ui,j } is an array of independent variables with identical distributions. In the laterally symmetric case, the above model (10) becomes:

$$X_{i,j} = \beta_1(X_{i-1,j} + X_{i+1,j}) + \beta_2(X_{i,j-1} + X_{i,j+1}) + U_{i,j} \quad . \qquad (11)$$

Ord (1975), following the joint probability of Whitte's definition, gave a more generalized definition of the term 'neighbour' which can be applied both to regular and irregular lattices and presented a general spatial autoregressive model as:

$$X_i = \alpha + \rho \sum_{j \in J(i)} w_{ij} x_j + \varepsilon_i \quad . \qquad (12)$$

where

$J_{(I)} = \{i_1, i_2, i_3, \ldots, i_{mi}\}$ whose variates interact with x_I

m_i the number of variates that interact with x_I

w_{ij} is the weight (or degree of association) between observations at site i and j.

and ε_I are independent error terms.

Ali (1979) discussed a generalization to the Whittle's two dimensional spatial model to spatial-temporal model. Let $X_{i,j}$ be as before, and for a discrete time t, denote the random variable by $X_{i,j,t}$. The following model is given by Ali as the generalization of Whittle's two dimensional spatial model to spatial –temporal model.

$$X_{i,j,t} = \sum \Phi_{r,c,k} X_{i-r,j-c,t-k} + \varepsilon_{i,j,t} \qquad (13)$$

where the sum is over all integer values of i,j,k, except for i = j = k = 0.

Apart from the GLS method, Ord (1975) discusses different practical estimation methods for the parameters of various spatial models. Many authors like Bartlett (1975), Cliff and Ord (1981), Bennet and Hanning (1985), Crissie (1991) give different spatial statistical models to represent the stochastic process which generates the particular spatial data and presented estimation procedures with examples.

APPENDIX 2

Data on grain yields of wheat for row planted fields (a cell represents 1 m^2).

RA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
	454	339	367	420	331	284	269	160	444	482	395	334	201	259	337	329	517	416	407	312	395	460	512	455	371	410	433	428	499	248	311	415	354	298	399	442	354	306	369	391
RB	379	292	342	408	391	369	298	220	441	423	358	304	404	301	311	311	362	361	288	322	359	382	389	308	364	386	343	393	478	297	286	384	350	344	223	425	370	394	475	423
RC	375	314	361	321	290	330	247	229	453	329	388	458	282	403	352	361	478	355	334	312	381	403	338	310	290	325	370	286	383	286	364	327	194	234	313	308	292	256	234	270
RD	298	364	338	271	307	195	227	207	365	348	250	202	272	245	266	173	288	240	246	188	322	316	325	346	288	307	385	371	420	275	384	441	332	278	363	307	285	375	374	362
RE	379	383	415	348	301	279	272	218	382	351	255	242	273	336	263	290	310	335	316	218	281	296	263	337	257	256	289	417	307	235	411	303	235	159	221	445	323	255	308	340
RF	429	277	300	264	235	338	337	353	342	370	269	335	362	324	366	350	457	396	326	288	347	282	354	371	308	344	350	305	283	390	304	404	372	247	292	494	342	417	403	388
RG	369	333	315	369	282	284	396	329	409	390	312	230	321	240	336	335	442	368	188	236	343	317	419	413	374	341	382	370	398	382	405	331	354	311	300	421	295	388	340	442
RH	420	384	335	308	334	275	337	228	304	335	198	110	291	331	191	227	329	397	405	282	311	336	359	423	317	327	374	329	439	325	358	326	295	258	336	416	285	379	290	264
RI	357	308	310	354	347	373	448	369	424	371	141	106	265	312	194	278	369	282	239	214	300	327	298	327	303	365	397	317	425	297	381	299	398	261	333	468	375	338	307	474
RJ	390	305	299	277	267	418	245	300	406	428	206	136	347	281	292	326	429	429	277	270	185	298	341	480	356	300	255	343	358	321	292	383	312	354	313	424	384	327	382	384
RK	373	327	336	326	274	431	336	336	292	430	151	150	258	384	259	316	323	359	261	251	225	240	249	298	241	263	355	376	316	343	395	496	483	368	365	289	302	311	396	387
RL	369	287	385	331	308	341	235	274	261	318	116	159	406	257	182	360	305	389	320	302	311	291	327	282	237	258	242	283	224	311	337	307	363	458	345	262	255	357	334	219

Grain yield row planted ctd.

	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
RA	344	136	214	282	355	299	294	395	344	255	313	409	275	369	408	359	334	328	321	324	315	294	301	404	329	432	300	317	243	331	317	350	322	254	325	482	302	341	327
RB	168	218	185	387	333	289	329	475	354	343	343	363	334	377	321	397	439	387	444	374	318	310	364	404	318	411	336	313	237	277	328	314	258	322	330	291	329	402	344
RC	201	245	317	261	210	359	276	329	249	281	296	314	275	286	317	308	303	267	291	333	317	294	270	314	283	351	309	316	328	328	259	330	320	410	351	420	383	353	350
RD	301	281	279	375	230	380	256	414	369	269	278	361	275	346	425	329	309	288	335	383	272	288	244	262	257	308	251	323	291	320	315	354	308	415	333	390	302	352	287
RE	257	215	225	287	146	315	255	323	262	255	256	354	264	298	270	244	238	247	266	248	256	287	302	281	230	262	275	260	227	321	350	322	292	318	238	281	324	384	371
RF	357	353	364	469	246	320	347	402	337	282	272	342	391	315	357	336	364	330	326	361	388	378	426	291	325	392	331	329	243	357	350	327	350	273	244	242	254	318	297
RG	334	343	383	347	288	453	447	354	363	419	360	435	369	399	395	424	389	423	324	365	410	330	334	396	321	292	259	298	276	278	460	354	407	324	328	318	320	316	309
RH	256	340	385	402	243	270	347	275	320	417	308	326	328	345	411	404	352	410	422	374	355	301	318	334	367	422	301	345	241	367	335	357	384	402	228	250	274	303	331
RI	316	276	364	397	285	338	376	380	377	403	393	341	342	342	369	374	279	329	425	244	286	269	217	242	275	308	314	290	173	270	339	333	233	284	191	411	371	389	321
RJ	338	399	298	312	232	279	319	320	330	365	338	257	323	310	250	362	312	365	333	256	241	249	291	271	281	254	316	325	258	341	296	264	358	341	128	255	249	334	262
RK	271	190	294	270	218	301	433	305	280	345	306	314	334	362	356	340	335	385	388	484	364	373	285	329	292	390	321	259	172	336	312	247	393	377	139	224	368	396	351
RL	304	294	222	303	171	153	335	307	294	330	266	318	211	308	266	272	226	274	269	388	281	249	145	181	168	123	213	230	242	236	261	185	327	324	66	139	319	453	387

APPENDIX 3

Data on grain yields of wheat for broadcast sown fields (a cell represents 1 m^2)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
BA	207	257	219	126	141	222	97	123	219	210	195	269	243	445	285	164	145	176	244	221	164	177	273	179	295	263	296	258	387	632
BB	200	180	154	81	101	142	109	110	170	261	170	189	154	243	148	105	136	141	223	192	134	163	114	209	171	148	144	216	240	222
BC	301	217	245	201	189	143	100	136	171	323	305	202	120	231	284	211	182	239	246	215	313	290	227	213	148	305	233	197	200	215
BD	208	272	247	151	207	209	124	254	311	184	225	326	166	300	203	145	201	179	257	359	145	179	274	257	200	209	262	254	310	284
BE	202	173	209	200	235	140	180	214	255	206	229	143	208	266	179	132	153	105	231	266	150	135	217	150	128	141	312	228	250	187
BF	177	101	183	145	216	166	92	118	163	138	211	241	300	146	141	117	162	159	224	185	159	161	219	213	164	193	288	237	265	232
BG	156	154	161	113	132	108	180	184	236	108	184	259	159	158	162	146	118	143	200	196	154	112	208	233	162	117	110	186	204	134
BH	205	143	196	118	135	187	159	223	316	151	117	122	136	172	233	103	80	148	232	222	107	145	188	160	169	109	241	290	275	339
BI	198	227	247	273	168	191	155	186	324	235	255	216	156	237	139	245	198	272	253	299	190	140	165	275	184	273	289	226	311	284
BJ	148	154	279	226	160	177	239	275	282	260	152	287	231	314	188	191	222	338	374	218	255	298	318	323	245	216	230	111	279	257
BK	199	193	250	178	190	165	237	224	292	276	307	269	178	222	219	207	222	257	285	229	207	256	327	354	311	313	241	280	310	311
BL	253	163	156	124	156	250	89	130	189	205	197	165	76	206	178	161	122	156	177	217	212	301	260	305	276	255	333	164	286	249

Grain yield broadcast sown ctd.

	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
BA	235	308	315	344	333	363	355	283	295	309	366	325	296	497	386	397	411	449	321	413	325	353	331	315	389	284	382	294	360	290
BB	208	187	192	216	167	253	238	232	171	226	272	266	296	290	427	321	296	293	282	262	341	344	339	266	346	264	350	355	296	300
BC	317	253	243	152	177	123	195	243	126	220	247	152	130	307	371	312	281	386	327	244	304	243	296	288	292	246	360	258	331	321
BD	307	276	254	251	165	313	190	219	263	228	226	181	176	313	365	350	218	249	241	296	328	285	296	240	371	369	380	306	182	227
BE	275	281	201	195	201	135	220	208	233	231	168	195	212	245	382	355	184	278	247	268	360	326	217	151	277	230	351	379	247	277
BF	241	271	260	192	156	136	195	236	226	258	216	188	255	227	284	210	194	372	324	340	304	298	246	241	251	272	403	403	338	345
BG	235	256	226	227	370	220	248	420	314	319	309	278	382	348	411	341	299	352	259	289	390	347	375	296	242	208	277	235	265	250
BH	328	261	178	166	232	259	233	219	227	251	209	222	286	237	204	351	362	323	355	265	350	348	318	275	262	266	415	365	336	376
BI	289	257	235	258	271	289	235	238	290	280	232	222	246	173	181	259	371	586	308	401	314	307	374	227	262	239	311	273	240	326
BJ	249	338	188	205	168	264	217	181	309	294	268	301	311	171	194	272	268	275	370	273	290	274	360	324	259	331	249	177	186	287
BK	343	298	225	307	276	288	279	264	292	386	432	264	335	306	309	356	324	300	303	216	376	302	336	349	424	365	359	220	250	295
BL	163	214	230	248	255	282	308	270	332	283	284	299	324	239	281	320	238	360	242	230	330	387	383	271	364	251	304	322	348	310

REFERENCES

Ali, M.M. (1979): "Analysis of stationary spatial temporal processes: Estimation and prediction", Biometrika, 66, 513-518.

Bartlett, M. S. (1955): An introduction to stochastic processes. Cambridge University Press, Cambridge.

Bartlett, M. S. (1967): "Inferences and stochastic processes", Journal of the Royal Statistical Society 130, 457-478.

Bartlett, M. S. (1975): The statistical analysis of spatial pattern. Chapman and Hall, London.

Bennet, R. J. and Hanning, R. P. (1985): "Spatial structure and spatial interaction: Modeling approaches to the statistical analysis of geographical data", Jornal of the Royal Statistical Society 148, 1-127.

Besag, J. E. (1974): "Spatial interaction and the statistical analysis of lattice systems", Journal of the Royal Statistical Society 36, 192-225.

Besag, J. E. and Kempton, R. A. (1986): "Statistical analysis on field experiments using neighboring plots", Biometrics 42, 231-251.

Bonham, C. D., Reich, R. M., and Leader, K. K. (1995): "Spatial cross-correlation of bouteloua ragcilis with site factors", Grassland –science 41, 196-201.

Cahn, M. D., Hummel, J. W. and Brouer, B. H. (1994): "Spatial analysis of fertility for site-specific crop management", Soil Science Society of the American Journal 18, 1240-1248.

Chou, Y. H. (1992), "Spatial autocorrelation analysis and weighting functions in the distribution of wildland fires", International Journal of Wildland Fire 4, 169-176.

Cliff, A. D. and Ord, J.K. (1973): Spatial autocorrelation. Pion, London.

Cliff, A. D. and Ord, J. K. (1981): Spatial processes: Modeling and application. Pion, London.

Cochran, W. G. and Cox, G. M. (1957): Experimental designs. Wiley, New York.

Cressie, N. (1991): Statistics for spatial data. Whiley, New York.

Czaplewski, R. L., Reich R. M. and Bechtold, W. A. (1994): "Spatial autocorrelation in growth of undisturbed natural pine stands across Georgia", Forest Science 2, 314-328.

Dyke, G. V. and Grundy, G. M. F. (1988): Comparative experiments with field crops. Charles Griffin & Company, London.

Galton F. (1889): "Comments on Taylor, E. B. 'on a method of investigating the development of institutions, applied to laws of marriage and decent", Journal of the Royal Anthropological Institute of Great Britain and Ireland 18, 245-271.

Geary, R. C. (1954): "The contiguity ratio and statistical mapping", The Incorporrated Statistician 5, 115-145.

Girma, T. and Amsale, T (1999): "Results of uniformity trials for wheat experimental", Journal of the Ethiopian Statistical Association 9, 55-68.

Gomez, K. A. And Gomez, A. A. (1984):Statistical procedure for agricultural research. Wiley, New York.

Gupta, S. C,; Kauchik, L. S. and Yadava, T. P. (1991): "A study of size and shape of experimental plots with groundnut (Arachis hypogaea L.)", Indian Journal of Agricultural Research 25, 161-166.

Hathway, W. H. (1961): "Convenient plot size", Agronomy Journal 3, 279-280.

Hathway, W. H. and Williams, E. J. (1985): "Efficient estimation of the relationship between plot size and the variability of crop yield", Biometrics 14, 207-222.

Heplle, L. W. (1978): "The econometric specification of spatio-temporal models" in timing space and spacing time 3, editors Carlstein, T.; Parkes, D. Thrift, N. (Edward Amod, London), 66-80.

Katayal, V. and Rajput, R. K. (1978) : "Effect of size and shape of plots and blocks on variabilty in the yield of rice grown on saline sodic soils ", Indian Journal of Agricultural Sciences 48, 614-618.

Mead, R. (1988): The design of experiments: statistical principles for practical researches. Cambridge University Press, Cambridge..

Miron, J. (1984): "Spatial autocorrelation in regression analysis: a beginner's guide", in Calinski, T. and Klonecki, Whilmot eds.: Linear Statistical Inference, Proceedings of Pozonan Conference. Springer Verlag, Berlin , 200-222.

Mohammed, A. A. G. (1987): "Estimate of optimum plot size and shape for dura sorghum (sorghum bicolar) yield experiments", East African Agricultural Journal 52, 75-81.

Moran, P. A. P. (1950): "Notes on continuos stochastic phenomena", Biometrika 37, 17-23.

Narayana, M. R. and Ramanata, C. K. (1983): Modified use of Smith's law for coefficient of heterogeneity on the basis of field experiments conducted on the size and shape of plots and blocks for dryland castor", Indian Journal of Agricultural Sciences 53, 250-255.

Narayana, M. R. and Ramanata, C. K. (1985): "Experimental plot size and shape based on data from uniformity trial with dryland pigeonpea", Indian ournal of Agricultural Sciences 55, 538-543.

Ord, J. K. (1975): "Estimation methods for models of spatial interaction", Journal of the American Statistical Association 70, 120-126.

Ostle, B. and Mensing, R. W. (1975): Statistics in research. Iowa State University Press, Iowa.

Papadakis, J. S. (1937). Agricultural Research: Principles, Methodology, Suggestions. Edicion Argentina, Buenos Aires.

Ram, B.; Agarwal, M. C.; Gupta, R. K. and Babu, R. (1993): "Size shape and arrangement of plots for field trials on different rangelands of India", Range Management and Agroforestry 14, 117-124.

Smith, H. F. (1938): "An empirical law describing heterogeneity in the yields of agricultural crops", Journal of the Agricultural Science 28, 1-23.

Stephan F. F. (1934): "Sampling errors and the interpretation of social data ordered in time and space", Journal of the American Statistical Association 29, 165-166.

Sukal, R. R. and Wortenberg, D. E. (1983):"A test of spatial autocorrelation analysis using an isolation –by –distance model", Genetics 105, 219-237.

Tobler, R. R. (1970): "A computer movie simulating urban growth in the Detroit region", Economic geography supplement 46, 234-240.

Whittle, P. (1963): "Stocastic process in several dimensions", Bulletin of the International Statistical Institute 40, 974-994.

Zhang, W. and Cho, W. (1995): "Spatial autocorrelation analysis of trace elements of soil in Tianjin plain area", Sinica 32, 50-57.

DECLARATION

This thesis is my original work and has not been presented for a degree in any other university

Name Kinfemichael Alemu

Signature _____

This thesis has been submitted for examination with my approval as University Advisor

_____

DR. Butte Gotu

June, 18.00

Date