THE PROFITABILITY OF PRECISION AGRICULTURE IN THE BOTHAVILLE DISTRICT

Ntsikane Maine

Submitted in accordance with the requirements for the degree

Philosophae Doctor

in the
Faculty of Natural and Agricultural Sciences
Department of Agricultural Economics
(Centre for Agricultural Management)
University of the Free State

Promotors:

Dr W.T. Nell
(University of the Free State)

Prof. J. Lowenberg-DeBoer
(Purdue University)

Co-promotor:
Dr Z. Gudeta

Bloemfontein
November 2006

Acknowledgements

During the course of undertaking the research and writing up this thesis, I received support, guidance, assistance and prayers from a number of people. I would like to thank all those who helped me during my entire study and made a success out of it.

Firstly, I extend my sincere thanks to my internal promoter, Dr Wimpie Nell; for his guidance, advice, and devotion towards this study.

A special word of appreciation to my external promoter, Prof. Jess Lowenberg-DeBoer of Purdue University, for all his technical inputs and prompt attention to my enquiries during the course of my study. His understanding and patience meant a lot.

I attribute special acknowledgements to my co-promoter, Dr Zerihun Gudeta, for all his support and valuable inputs.

My appreciation is extended to Mr Thabo van Zyl, owner of Rietgat farm, for allowing me to conduct this research on his farm, for the information he provided, and for always being enthusiastic to my calls.

Thank you very much to Dr Charles Barker, a GIS specialist at the University of the Free State, for his technical advice on GIS and help in aggregating data.

I express my gratitude to Mr Cobus Burgers and Ms Adri Meistri of Omnia, as well as Dr Herman Booysen of NetGroup for digitising maps and helping in collating data. Thank you for your time and the information you provided.

Special thanks to Mr W.E. Oosthuysen of Suidwes Agriculture, Dr P. LeRoux and Dr G. Ceronio of the Department of Plant and Soil Sciences for all the technical information on precision agriculture and soil data used in this thesis.

Technical assistance of Senwes in grid soil sampling of the research field is acknowledged.

The inputs of the individuals on the ListServe of GeoDa™ statistical analysis are highly appreciated, particularly the help of Julia Koschinsky and Terry Griffin.

I thank the University of the Free State and the Faculty of Natural and Agricultural Sciences for giving me the opportunity to study at this institution. A word of thanks to the staff of the Department of Agricultural Economics and the Centre for Agricultural Management, particularly Mrs Dora Du Plessis for editing and attending to the technical aspects of this thesis.

I am also indebted to numerous friends and family for their moral support, prayers and encouragement. Thank you all. I thank my late parents, to whom I dedicate this piece of work.

To my husband, Teboho, thank you for the emotional support and putting up with my odd working hours.

The financial assistance of New Holland South Africa and the Andrew Mellon Foundation is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily attributed to either organisation.

Finally, I acknowledge and thank my heavenly Father for giving me strength and courage, and making everything possible. I thank you, Lord.

Summary

Variable-rate application technology based on soil potential and other field attributes is gradually replacing the standard rates of fertilizer application for individual cropping systems. In South Africa, differential application of inputs in cash crop production is mainly concerned with fertilizer and lime, and this indicates the importance of these inputs. This study evaluates the maize yield response to variable-rate (VR) application of nitrogen (N), and estimates the profitability of VR application of N relative to single-rate (SR) application under South African conditions.

Data was collected from an experimental field of 104 ha on a farm in the Bothaville district. A strip-plot design consisting of 180 strips was used for this on-farm research experiment. This design involved treatments that ran in the same direction across the field as planting and harvesting. The objectives were to determine the maize crop response functions under different N rates, to estimate optimal N rates for different management zones in different years, and to assess profit estimates using ordinary least squares (OLS) and spatial error (SER) models. The methodology involves modelling maize yield response functions for N. A Baseline regression model that analyses variable-rate technology as a package was used, while three sensitivity tests were used to determine the consistency of the estimates.

The results of this study indicate that there is a significant variation in maize yield response to the applied N on the basis of the application method used. Profit analysis resulting from the application strategies indicates that, in general, VR results in higher farming profits than SR. The analysis indicates that yield obtained from VR strategy can compensate additional costs incurred with the investment in VR technology. This finding is consistent in all the models. It has been established that yield response to fertilizer depends on soil conditions such as the effective soil depth, which has a positive effect on yield. Yield response also differs among management zones.

Differences were observed between the results obtained from the OLS models and the results obtained with the SER models, and this has an impact on decision-making. The importance of taking spatial effects into account came to the fore, as inaccurate results can be obtained with methodologies that ignore the spatial dependencies in the analysis of yield monitor data.

Key terms: Precision agriculture, precision farming, variable-rate application, single rate application, profitability, spatial error model, nitrogen, South Africa.

Opsoming

Veranderlike-toedieningspeiltegnologie, gebaseer op grondpotensiaal en ander landeienskappe, is besig om die standaardpeile van kunsmistoediening vir individuele gewasstelsels geleidelik te vervang. Differensiële toediening van kontantgewasproduksie-insette fokus in hoofsaak op kunsmis en kalk. Hierdie studie evalueer die mielieopbrengs-reaksie op veranderlike toedieningspeile (VT) van stikstof (N), en beraam die winsgewendheid van VT van N in vergelyking met enkelpeiltoedienings (ET) onder Suid-Afrikaanse toestande.

Die data is afkomstig van 'n eksperimentele land van 104 ha op 'n plaas in die Bothaville-distrik. 'n Strookperseelontwerp, bestaande uit 180 stroke, wat in dieselfde rigting as die plant en stooprigting uitgelê is, is gebruik. Die doelstellings was die bepaling van mieliegewas-responsfunksies aan die hand van verskillende N-peile, die beraming van optimale N-peile vir verskillende bestuursones in verskillende jare, en die analisering van boerdery winsskattings met die gebruik van gewone kleinste kwadrate-modelle (GKK-modelle) en ruimtelike fout-modelle (RF-modelle). Die metodologie behels die modellering van mielieopbrengs-responsfunksies vir N. 'n Basislynregressiemodel, wat VT-tegnologie as 'n pakket ontleed, is gebruik, terwyl drie sensitiwiteitstoetse gebruik is om die konsekwentheid van die beramings te bepaal.

Die resultate van hierdie studie dui aan dat daar beduidende variasie bestaan ten opsigte van mielieopbrengsreaksie op die toedieningspeile van N, afhangende van die toedieningsmetode (veranderlik of enkel) wat gebruik is. Die winsontleding wat voortspruit uit die toedieningstrategieë toon aan dat VT oor die algemeen hoër winste as ET genereer. Die analise toon aan dat die opbrengs van 'n VT-strategie kan vergoed vir die bykomende koste wat met diebelegging VT tegnologie gepaard gaan. Hierdie bevinding is konsekwent in alle modelle. Daar is verder vasgestel dat die opbrengsrespons op kunsmis afhang van

grondtoestande, soos die effektiewe gronddiepte, wat 'n positiewe uitwerking op opbrengs het. Verskillende bestuursones het ook verskillende opbrengsreaksies getoon.

Verskille wat waargeneem is tussen die resultate wat met die GKK-modelle behaal is en die resultate wat met die RF-modelle verkry is, het 'n impak op besluitneming gehad. Die noodsaaklikheid om ruimtelike effekte in ag te neem, is beklemtoon aangesien onakkurate resultate behaal kan word met metodologieë wat die ruimtelike afhanklikhede in die ontleding van opbrengsmonitordata ignoreer.

Sleutelterme: Presisielandbou, presisieboerdery, veranderlike-toedieningspeiltegnologie, enkelpeiltoedienings, winsgewendheid, ruimtefout-model, stikstof, Suid-Afrika.

Table of Contents

List of Tables

Table

List of Figures

Figure

Chapter 1

INTRODUCTION

1.1 BACKGROUND

Fertilizer is one of the most important inputs in crop production throughout the world. Simpson (1986) reports evidence obtained from many years of experimentation that nitrogen (N), together with phosphorus (P) and potassium (K), have the greatest effects of all fertilizer nutrients on crop yields. Consequently, these elements are supplied in easily absorbable forms. As Deckard, Tsai and Tucker (1984) point out, fertilizer N is used increasingly to supplement soil N for the production of food, animal feed and fibre for an ever-increasing world population. Although the usage of N and P continues to increase steadily worldwide, usage patterns within individual countries vary greatly according to factors such as the wealth of the country concerned, the population growth rate, the development of agricultural technology, the extent of concern over water purity, and environmental damage (Parkinson, 1993).

The optimal crop requirements for fertilizer nutrients vary widely from one year to the next, and this same variation also occurs within the same field due to differences in soil types. It is therefore important to determine the precision (site-specific) applications of fertilizer that adequately meet the crop needs and reduce environmental degradation without yield loss. The balance between nutrient supply and demand is one of the central themes of sustainable agriculture. In Europe, a principle of balanced fertilization is followed, according to which fertilizer recommendations are based on the foreseeable nutrient demands and the soil nutrient supply. The legislation differs from country to country, but in general there is increasing worldwide concern about protection of the environment (Vermeulen, Steen & Schnug, 1998).

According to Henao and Baanante (1999), more N and K than P are depleted from African soils, primarily due to leaching and soil erosion. The majority of these soils are classified as marginal, and are considered unsuitable for crop production. Although soil development projects have been undertaken in South Africa to a certain extent, some soils are marginal and degradation and soil erosion are evident (Department of Environmental Affairs and Tourism, 1999). As soils in South Africa are generally thin and their fertility is moderate, an attempt to supplement N- and P-deficient soils led to an increase in fertilizer use in the 1960s and 1970s. However, this practice had a negative impact since it led to an increase in soil acidity (Morrison & Pearce, 2000).

1.2 PROBLEM STATEMENT

The increased use of fertilizer in South Africa creates its own problems. A vigorous annual increase in the use of N, P, and K fertilizers has been observed in South Africa and throughout the rest of the world, particularly N fertilizers in areas with more developed agricultural systems (Simpson, 1986). This is the result of the effort to intensify production to maximise yields, leading to inefficient and uneconomic use of fertilizer nutrients and the contamination of watercourses by leached nutrients, especially nitrates. Between 40 and 70% of N applied to crops in the form of fertilizer is taken up by the first crop after fertilization, and most of the portion that is not absorbed is lost by leaching, denitrification or volatisation**.** It is important to assess the viability of implementing the variable rate application (precision application) of fertilizer according to soil yield potential, in an effort to reduce the loss of N and improve the efficiency of fertilizer use.

There is concern worldwide about the increasing concentrations of nitrates in surface water, ground water, lakes and the marine environment. This increase is unlikely to be the result of a single environmental factor, but rather of a composite of factors. Much of the blame has been directed towards the intensification of agricultural production, as this leads to N enrichment. Although this may be true, other factors such as increased N flux from the atmosphere to the terrestrial environment and an increase in the N loading from human (sewage) sources compound the problem (Burt, Heathwaite & Trudgill, 1993).

According to Burt *et al.* (1993), there is strong evidence to suggest that, over the period 1973 to 1993, the nitrate issue escalated from just a local pollution problem in developed countries to a worldwide problem. Unless measures are taken to reduce the rate and magnitude of nitrate input into receiving waters, the nitrate issue will ultimately have adverse effects, especially on the human health. In consideration of the perception that agriculture is the main source of pollution in aquatic systems, research that focuses on reducing nitrate and other fertilizer nutrient losses and the economies thereof, is therefore a priority.

Fertilizer use is influenced by many external factors. The price control on fertilizer that was lifted in 1984 had serious financial implications, both for farmers and the fertilizer industry in South Africa. Prior to 1984, all prices and imports of fertilizer were controlled (Fertilizer Society of South Africa, 1989). The cost of fertilizer has increased tremendously in recent years, forcing farmers to find means of using it more efficiently. Simpson (1986) points out that there are many ways in which fertilizer costs can be reduced without adversely affecting yields or efficiency. Some of these methods have no effect on farming operations and can be adopted immediately, while others such as variable-rate (VR) fertilizer application involve radical changes and/or high investment in technology, as well as additional management capacity.

VR application technology advice based on the soil potential and other field attributes is gradually replacing the standard rates of fertilizer application for individual cropping systems (Burt *et al.,* 1993). According to Matela (2001), differential application of inputs in South African cash crop production is mainly concerned with fertilizer and lime, which indicates the importance of these inputs for cash crop production in South Africa. Burt *et al.* (1993) argue that the change in the form and application method of fertilizer, especially N, is mainly in response to the changing price per unit of N, rather than considerations regarding the likely efficiency of use. In contrast, Matela (2001) found that farmers in South Africa adopt VR technologies to improve efficiency, which leads to reduced per unit cost of production, and ultimately a probable increase in farming profit.

Nutrient management has been variable (site-specific) for a number of years, and the technology has changed progressively to make efficient precision (VR) management of nutrients possible. The one thing that has not changed over the years and remains the concern expressed most frequently by producers, is the profitability of the precision management of nutrients (Malzer *et al.*, 1999), and more specifically, at the present stage, the profitability of the VR application of inputs such as seed, nitrogen, phosphates and lime.

Precision agriculture is often described as the imminent revolution in agriculture. However, with the exception of South Africa, precision agriculture is only practised to a limited extent in Sub-Saharan Africa (Nell, Maine & Basson, 2006) – despite the remarkable capabilities and adjustment of the technologies developed to support this approach to farming. Key issues concerning the profitability and environmental consequences of this technology remain largely unresolved (Lowenberg-DeBoer & Swinton, 1997; Malzer *et al.*, 1999; Weiss, 1996). According to Lowenberg-DeBoer (1999a), the actual investment in precision farming in some parts of the United States of America (USA) has been promising, but it is considerably more modest than what is portrayed by the media. Precision agriculture is an intuitively appealing concept, but the profitability of the practice remains uncertain, and this affects its widespread use (Lowenberg-DeBoer, 1999a).

Malzer *et al.* (1999) identify the real challenge to precision agriculture as determining the factors or items that influence crop production for a given field, and developing an appropriate strategy to maximise profitability for the producer. In a study carried out by Matela (2001), farmers cited the potential increase in profitability as one of their main considerations in adopting precision agriculture technology. In order to promote adoption of this technology, relevant research needs to be conducted so that information about its benefits can be provided to farmers, agri-business and society at large.

1.3 OBJECTIVES

The main objective of this research is to evaluate the profitability of precision agriculture, as applied to a maize field in the Bothaville district of the Free State Province in South Africa. Profitability as described by Van Zyl *et al.* (1999) means positive net returns on the investment. Precision agriculture is defined as a management strategy that uses information technology to bring data from multiple sources to bear in decision-making (National Research Council [NRC], 1997).

1.3.1 Sub-objectives

The sub-objectives are to:

- determine the response of maize yields to N in different management zones;
- estimate optimal N rates for different management zones in different years;
- estimate optimal N rates for both VR and single rate (SR) treatments in different years;
- compare the profitability of VR application of N with the profitability of SR application;
- assess profit estimates using ordinary least squares (OLS) and spatial error (SER) models.

NB: Maize yields are measured in metric tons per hectare, whereby 1 000 kg = 1 ton.

1.3.2 Hypotheses

1. There is no auto-correlation in yield monitor data obtained from the study field.
2. Maize yield response to N varies by management zones; there is also variation in yield obtained from different management zones.
3. There is a statistical difference between the profitability of VR and SR applications of N.

1.4 METHODOLOGY

Brouder and Nielsen (2000) recommend following a systematic approach that is common to all research projects in order to conduct a successful on-farm trial. This involves formulating a hypothesis or research question, planning an experiment or trial to objectively test the question or hypothesis (experimental design), careful observation and collection of data from the experiment, and interpreting the experimental results to answer the research question (accept or reject the null hypothesis).

Experimental research was carried out as part of this study. This involved trials on a farm named Rietgat, situated in the Bothaville district of the Free State Province. Although soils in Bothaville are generally perceived as homogeneous, yields in this area show great variation. Such soils could therefore provide an interesting analysis of precision agriculture with regard to the profitability of VR application of N relative to SR application.

A strip-plot design, as recommended by Brouder and Nielsen (2000), was implemented. A 104-hectare field was divided equally into strips, with alternating sets of six rows for VR application and six rows for SR application of N. Each set of six rows (strip) constituted a block (plot). This design is used when it is known for certain that differences between the experimental units (soils) exist in such a way that they contribute to the variation between the tested factors (Safo-Kantanka, 1994). A full discussion of the methodology followed in this study is provided in Chapter 3.

1.5 VALUE OF THE STUDY

Fertile soils are high on the list of the most important resources on earth, as they are the source of life for future generations. As Schnug, Panten and Haneklaus (1998) put it: "Sustainable agriculture should use these resources in such a way that the present and future human needs for food and other agricultural goods are warranted, whereas the quality of the environment and the natural resources remain preserved." Sustainability of agriculture was at the top of the agenda at the World Summit in Sustainable Development held in Johannesburg, South Africa, in 2002. The massive media coverage on this summit suggests that the majority of South Africans are conversant with the concept of agricultural sustainability, and that greater assistance is needed in conducting sustainable farming operations.

The improper use of nutrients in agricultural soils of South Africa contributes to environmental problems, and this problem needs to be addressed while maintaining the profitability of agriculture. Precision agriculture is a complementary process involving different types of best management practices, including environmental management. The view of Mausbach, Lytle and Spivey (1992) is that the philosophy of precision agriculture can also be regarded as a theory of environmentally sustainable farming, thus indicating the positive social impact of precision agriculture. Even though the quantification of environmental benefits is not the focus of this study, this aspect should not be neglected in determining the value of precision agriculture.

Precision agriculture can aid in the reduction of nutrient loss to the environment by applying the correct quantities to the appropriate areas. Most importantly, precision agriculture has the potential to increase the profit of farming operations in South Africa, which face a cost-price squeeze, so that the growth and sustainability of these farming operations can be realised. VR fertilizer applications include one of the valuable aspects of precision agriculture, as many research studies in the USA indicate a potential decrease in per-unit production cost after applying nutrients according to soil requirements (VR application) versus a general (SR) application on the entire field (Forcella, 1993; Lowenberg-Deboer & Boehlje, 1996; Olson, 1999). Malzer *et al.* (1999) conclude that VR management of fertilizer nutrients has the potential to substantially improve the economic return to producers by 10-20%. If precision agriculture can reduce costs and improve economic returns to producers under South African conditions too, it can go a long way towards increasing farming productivity and sustainability, and could even contribute to the growth of the country’s economy. The potential increased benefits may, however, also be associated with some risks, and the costs incurred by making the wrong decision may be substantial – a factor that is paramount to sustainable production, and should be taken into consideration (Forcella, 1993).

Since precision agriculture requires localised information regarding cause and effect relationships with seeding rates, fertilizers, and other agro-chemicals, on-farm experimentation is necessary to promote effective use of this new technology (Napier, 2001). The on-farm experimentation conducted for this study endeavoured to establish the relationships between yield as a dependent variable and different N rates (explanatory variable) under South African conditions.

A new focus on spatially detailed information obtained from precision agriculture presents an opportunity to agronomists, agricultural management specialists and agricultural economists to tackle problems associated with spatiality. The need to quantify the costs and benefits of detecting and exploiting spatial variability, particularly with regard to the soil characteristics and crop yield, is one of the most important problems addressed by this study. The resulting answers may have implications for farm management, and to some extent agricultural and environmental policy. This study illustrates the importance of taking spatial associations present in yield monitor data into account in order to provide accurate estimates. It is shown that spatial economic models are more accurate than traditional OLS models.

This research was collaboration between a farmer in Bothaville, researchers and extension services. This kind of collaboration between researchers and on-farm experimentation allows ongoing learning about new technologies, facilitates the development of decision-making models and establishes a basis for low-risk adoption of such models.

This research can contribute to the theory of agricultural management by providing greater insight into the value of precision agriculture, particularly VR technology as an agricultural management tool, with a view to enhancing the level of agricultural management in South Africa. The strategic approach to, and the risk involved in this new technology, can be identified in this study. According to Blackmore (1994), many different factors need to be considered in the formulation of a strategic approach, including balancing the potential economic returns with environmental impact and the degree of risk involved. As precision agriculture is capital-intensive and requires a large capital outlay, farmers have to be certain of the outcome of their decision to adopt this technology. It should be borne in mind that the interest of the farm operator lies in the economic returns on this technology, while agri-businesses – especially dealers – will be concerned about the sales that can be generated. However, this study adopted a neutral approach, and its sole purpose was to determine whether precision agriculture is profitable or not.

1.6 OUTLINE OF THE STUDY

This research experiment addresses the issue of the economic profitability of the VR application component of precision agriculture technology. The research problem and the background to the problem, the objectives and the motivation for the study are discussed in **Chapter 1**. Issues relating to soil nutrition and precision (site-specific) management of nutrients and crop response functions, as well as information on precision agriculture and the profitability analysis thereof in agronomic crops, are discussed in **Chapter 2**. The benefits of, and risks associated with precision agriculture, are also reviewed in Chapter 2. The conceptual framework and experimental design, procedures and techniques used in this research are presented in **Chapter 3**. Spatial econometrics (a relatively new statistical technique) is the main method of data analysis that is employed, and this is also explored in Chapter 3. The background and description of the study area, as well as a detailed discussion of the unit of analysis, are also attended to in Chapter 3. **Chapter 4** looks at the exploratory and descriptive analyses of the data. Regression and profitability analyses of the data are presented and discussed in **Chapter 5.** This report is concluded by **Chapter 6**, which contains a summary of the findings, conclusions and recommendations. The limitations of this study and the lessons learnt from it are also highlighted in Chapter 6.

Chapter 2

SOIL NUTRITION AND PRECISION AGRICULTURE

2.1 INTRODUCTION

Fertile soils are an important resource in agricultural production. It is pertinent that these resources are used in a sustainable manner so that provision can be made for present and future human needs for food and other agricultural products, while preserving the quality of the environment and the natural resources (Schnug *et al.*, 1998). Through the ages, agricultural production systems have benefited from the incorporation of technological advances, which were primarily developed for other industries. Mechanisation and synthesised fertilizers were introduced by the industrial age, while genetic engineering was offered in the technological age. Now the information age brings the potential for precision agriculture, one of the farming methods that may help to steer agriculture along a sustainable path.

Napier (2001) identifies precision agriculture as one of the technological developments that could precipitate a revolution in agriculture. Most of the new technologies require increased managerial capabilities and a larger scale of production for effective adoption. Large-scale production gives rise to increased productivity. This leads to a continued upward pressure on the supply of commodities and a downward pressure on prices – the well-known “treadmill effect” (Napier, 2001; Lowenberg-DeBoer & Swinton, 1997). On this treadmill, every opportunity should be taken to reduce costs and increase value. According to Napier (2001), one of the outstanding characteristics of leading farm managers is their ability to select and manage new technologies. Successful farm managers work closely with researchers to gain experience in managing new technologies, and to be the first to implement the latest technology. Furthermore, researchers can provide assistance with production decisions and facilitate the implementation of advanced techniques for crop production and soil fertility management.

Although precision agriculture is applied in South Africa, it is unknown in the rest of Sub-Saharan Africa. Considering the problems faced by Sub-Saharan countries (soil nutrient depletion, soil erosion, inadequate use of fertilizer and changes in production practices), precision agriculture could play a vital role in these countries if adapted to farmers' current production practices and level of education. Precision agriculture facilitates spatial management of land and optimal use of scarce agricultural inputs (seed, fertilizer and pesticides) available to farmers in developing countries. Technologies adopted should be tailored to meet a specific country's production constraints (Gandonou *et al.*, 2004). In the following sections, the effect of precision agriculture on soil nutrition and crop responses to applied nutrients will be explored.

2.2 SOIL NUTRITION

Management of soil nutrition is one of the largest input costs in the cultivation of crops in South Africa, especially maize (AGR318, 1998), and precision agriculture promotes efficient application of fertilizer to minimise costs and optimise productivity. In order to ensure a good harvest, soil nutrition should be adequate to promote optimal plant growth. Furthermore, these nutrients must be available in the right quantities, be applied at the right time and in forms absorbable by plants. In any fertilization programme, nutrient balances, nutrient use (especially nitrogen) and the effects of nutrients on crop yield and quality must be taken into consideration. Correct amounts of nutrients can be applied in the right places with precision agriculture technology, and easy computation of nutrient use and nutrient balances can be facilitated.

2.2.1 Nutrient balances

The nutrient balance, i.e. the balance between the supply and removal of plant nutrients, is one of the central themes of sustainable agriculture (Godwin *et al.*, 2002). Vermeulen *et al.* (1998) expand on the balance concept by stating that the annual difference between the total quantity of nutrient inputs and outputs can be calculated through the surface balance, which can indicate either a surplus or a deficit of soil nutrients. In the case of nitrogen (N), the difference between input and output lies in the N losses to air and water, as well as the N immobilised in the soil. The difference is not all-inclusive as certain processes, which are usually difficult to measure but contribute towards the surplus, are not taken into account. Mineralisation also takes place in the soil, adding positively to the surplus. These factors must be taken into account when determining N application rates.

In most African countries, a net loss of soil nutrients is usually observed because fertilization is practiced to a limited degree, and – in some cases – not at all (Wallace & Knausenburger, 1997). However, in the agricultural systems of industrialised countries, nutrient balances usually indicate a nutrient surplus. According to Vermeulen *et al.* (1998), this implies that more nutrients are applied to the soil than are removed. The implication is that nutrients not used by plants can be regarded as a waste, and increase the costs to the producer. Precision application of nutrients enhances nutrient balance by ensuring that specific amounts of nutrients are applied where necessary. Godwin *et al.* (2002) observed that simple N balance calculations have indicated that, in addition to a modest increase in yield, N surplus can be reduced by approximately one third with the spatially variable application of N. Fertilizer application is very important in crop production in terms of the associated costs and the impact on crop yield and quality.

2.2.2 Nutrient use and crop yields

In cash crops, the relationship between fertilizer use and crop yield has both positive and negative aspects. Each increment in fertilizer application results in a progressively smaller increase in yield, until a maximum yield is reached. Beyond the maximum, any increase in fertilizer rate will lead to no further increase in yield, or could even result in a decline. It is therefore uneconomical to apply more fertilizer beyond the maximum point.

Within phase II of the production function, there is a point below the maximum yield level, the optimal yield level, which corresponds to the optimal fertilizer rate. On the assumption that the objective of a farmer/producer is to maximise the expected profit, the necessary condition for profit maximisation is that, at the chosen fertilizer application rate, the input/output price ratio equals the slope of the yield response function. This is therefore the rate that maximises profit (Bullock *et al.*, 1998). According to Alivelu *et al.* (2003), economically optimal rates of fertilization can be calculated by equating the first derivatives of response equations to a selected fertilizer:crop price ratio, and solving for *x*. The optimal fertilizer application rates depend on input and output price relationships, as well as fixed costs for precision agriculture (Hurley, Malzer & Kilian, 2003). If product prices increase more rapidly than the fertilizer price, the most profitable production and fertilizer application level will be closer to the maximum production level. If input price increases at a higher rate than output price, input applications should be reduced, and may even stand at zero (Boehlje & Eidman, 1984; Van Zyl *et al.*, 1999).

Simpson (1986) advises farmers to aim to apply fertilizer nutrients at a rate close to the optimum. Simpson (1986) found that applying fertilizer at less than the optimal rate produces smaller yields, and usually results in a cash crop profit reduction. Insurance fertilization, i.e. applying more than the estimated optimum, is preferred by some farmers. However, reduced yield remains a potential penalty in this system, especially during drought spells.

Although the standard production theory will be used in this study, it is worth noting the applicability of the Von Liebig theory. Recommendations arising from the Von Liebig theory, also termed the law of the minimum, differ from those derived from the standard production theory. According to this theory, crop response can be determined by the most limiting (scarcest) production factor, until another factor becomes limiting. An increment of a non-limiting factor does not affect yield, assuming complementarity between the inputs. A plateau is reached when sufficient quantities of two inputs, x_1 (fertilizer) and x_2 (water), are added. However, once the plateau is reached, increasing either input does not change the output (Berck, Geoghegan & Stohs, 1998).

It is worth noting that the optimal fertilizer rates vary widely, depending on the type of crop, soil characteristics, the season and the area. Different application techniques, including precision agriculture, should take cognisance of this (Simpson, 1986). As optimal fertilizer rates differ from one soil type to the next, precision agriculture can allow variable application, which still meets the optimal application rates. In variable-rate application of fertilizer, the effect of fertilization on crop quality should also be taken into account.

2.2.3 Fertilizer use and crop quality

Fertilization affects the quality of most crops. Appropriate fertilization, particularly with N, can enhance the protein content of most grain crops to such an extent that it can exceed the quantity required for maximum grain yield production. The protein level of grains is particularly important in developing countries such as South Africa, where most of the human intake of protein is obtained from grain foods (Olson, 1984).

The response of yield and protein content to N supply also depends to a great extent on the crop cultivars and environmental conditions, especially the quantity of water and the times at which it is available to the crop. The latter is particularly important in rain-fed production under semi-arid conditions, as in this study. The increasing use of N does not always result in improved quality, and tradeoffs are sometimes made between quality and quantity.

The economic exploitation of spatial variability in N use is illustrated by the increasing use of precision farming in sugar beet cultivation (Weiss, 1996). Boosting the soil N levels tends to increase yield for this crop, while decreasing quality. This suggests that there is a profit-maximising N level, from which profit can be affected either upwards or downwards. The net profit changes resulting from these deviations are, however, not universal. In principle, precision farming can either increase or decrease the use of fertilizer. The spatially detailed information can be used to avoid fertilization in excess of plant requirements and reduce the possibility of run-off and leaching to ground water, and can also assist in economic analysis when estimating crop response functions (Weiss, 1996).

2.2.4 Crop response functions

Crop response functions can be useful in making decisions regarding the optimal input mix and the most profitable level of output. This kind of information is still required in precision farming, as in conventional farming, in order to make more accurate management decisions. Response curves are produced by varying the soil fertility and measuring the ensuing yield (Hancock, 2002).

Lowenberg-DeBoer and Boehlje (1996) identify an approach that has gained popularity in assessing the benefits of variable-rate N application, namely the estimation and comparison of site-specific crop response functions using multiple regression analysis. As supported by Heady and Dillon (1961), one reason for estimating agricultural production functions is the ability to provide basic scientific knowledge regarding input-output relationships. With knowledge of the appropriate relationships and economic principles, more practical recommendations and inferences can be made at field level. When used for economic analysis and recommendation, production functions provide one of the two sets of information needed for selection and decision-making. This information can be in the form of price data or other quantities, which serve as the basis for determining economic criteria. The physical quantities derived from production functions also constitute an essential part of the data required for input application decisions with regard to precision agriculture (Heady & Dillon, 1961).

Described in simple terms, physical production is a function of many resources. For instance, two resources, X and Z, are variables and a single product, Y, results. The crop response function, $y_i = f\ (x_i,\ z_i)$, therefore represents the yield in each sub-field (management zone); i = 1, 2, … *M,* is a function of soil fertility level *z* and applied input *x* with $f_x > 0$, $f_z > 0$, $f_{xx} < 0$ (Isik, Khanna & Winter-Nelson, 1999). The soil fertility levels depend on the nutrient content and clay percentage of soils, and the average soil fertility within the field is represented by $\bar{z}$ (Heady & Dillon, 1961; Hurley *et al.,* 2003).

Heady and Dillon (1961) show that representing output as a function of input categories can result in numerous types of production surfaces. In a three-dimensional form where resource magnitudes are measured along horizontal axes or on an input plane, and output is measured vertically or in product space, several sets of interrelated quantities can be derived, but only two are relevant to this study.

Firstly are the input-output relationships, which can be represented by curves. These curves express output in relation to variable input (X) when the input (Z) is held constant in various magnitudes. Input-output curves can be represented in two dimensions. The slopes of the individual input-output curves indicate the marginal products of the variable input. The marginal products can be used in the application of economic principles when specifying optimal resource use. According to Hurley *et al.* (2003), in the case of a farmer whose objective is to optimise yield, *x* should be chosen in such a way that the value of the marginal product of an input equals its marginal cost. At a given soil fertility test level, aeration/drainage and cropping history, Hancock (2002) also recommends an application that is not aimed at a maximum yield, but at an optimum yield. This is based on the fact that, at some point on the production curve, the cost of an additional unit of fertilizer will equal the value of the increase in yield.

A second relationship is represented by product isoquants, which indicate a specific output level relative to all possible combinations of the two inputs, or factors, which will produce a specified output. Isoquants can also be represented in two dimensions. The slope of each isoquant indicates the rate at which one resource substitutes for or replaces the other in order to maintain output at a specified level (Heady & Dillon, 1961). Production functions make provision for the algebraic and arithmetic expression of resource quantities and the specification of relationships. The algebraic input-output relationship can be specified for one resource by holding the other constant at a specific level. The isoquant equations can be derived from the production equation by expressing the input of one factor as a function of the output level and quantity of the other resource (Heady & Dillon, 1961).

These input-output relationships can be demonstrated by a case where the farmer is faced with the two distinct technology choices of conventional production practices (single-rate fertilizer application) and precision agriculture (variable-rate application), as discussed in Isik *et al*. (1999). Carryover effects are taken into account in Isik *et al*. (1999) in determining the optimal level of input. Lambert and Lowenberg-DeBoer (2000) also included carryover in their analysis, and concluded that higher returns are obtained from variable-rate technology (VRT) strategies for N and phosphorus (P) with carryover management than with conventional uniform strategies. The fertilizer carryover model was developed by Kennedy *et al.* (1973). The authors found that, with the assumed carryover coefficient of 0.4, the optimal N application increased slightly when the carryover coefficient was raised from 0.0 to 0.4 and the resultant net revenue for one production period was negligible. However, the consequence of carryover became important for optimal fertilizer application where several crops were obtained from the same plant in succession (Kennedy *et al*., 1973). Nevertheless, these carryover effects are beyond the scope of this study.

Simpson (1986) points out that yield response to added nutrients and the quantity of nutrients required to give optimum yield for a particular soil/climate vary considerably from one area to the next, and even from one field to the next. In the case of N, yield response depends on soil conditions (temperature, moisture, compaction, pH and clay percentage). Positive N-clay response is usually observed in soils with low clay content, with the response being generally poorer in soils with higher clay content. The response also depends on the supply of plant nutrients. There must be a balance between the required N and adequate amounts of other plant nutrients essential to optimal growth (Berry, 1990). The response of different crops to specific nutrients should also be taken into account.

Fertilizer response curves can also be generated for different management zones by varying fertilizer application per zone and measuring the resulting yields (Hancock, 2002). Varsa *et al.* (2000) evaluate VRT as a management tool for potassium (K) fertilization of corn (maize) and soybean, with varying ranges of depth to clay pan. The yield responses to fertilization across grid cells are compared. Of the nine study grids, only two show a positive correlation between applied K, trifoliate-leaf K, and soybean yield. The two responding cells apparently have the greatest depth to clay layer. Varsa *et al.* (2000) attribute this lack of response on the site to extreme drought conditions throughout the growing season, excessive crop stress, and residual K from the previous year's application. In such stressful crop conditions, it is important to determine what factors are most strongly correlated to yield.

On the whole, knowledge regarding input-output relationships can assist scientists in developing tools that farmers can use to determine the best crop management practices, and to suggest guidelines for recommendations regarding the situation of specific farmers. Under yield-optimising conditions, the extent of an input should be chosen in such a way that the value of the marginal product is equal to its marginal cost. Yield response to applied inputs and the input extent required for optimum yields for a particular soil/climate vary from one area to the next, from one field to the next and even within the same field, and this variability can be taken into account by application technologies such as precision agriculture.

2.3 PRECISION AGRICULTURE

As discussed in Chapter 1, the aim of this study is to evaluate the profitability of precision agriculture, and it is essential to present a clear view of this technology. This section reviews various themes related to precision agriculture. These include different precision agriculture (PA) techniques, information, profitability, risk, benefits and variability. Godwin *et al.* (2002) describe precision agriculture as a term defining a method of crop management that entails management of areas within a crop field that require different levels of input. Lowenberg-DeBoer and Boehlje (1996) define precision agriculture as monitoring and control applied to agriculture, including site-specific application of input, timing of operations and monitoring of crops and employees. In simpler terms, precision agriculture is therefore a way to help a farmer manage variation within his fields in a more proactive manner, by recognising site-specific differences in variables such as yield, soil texture, soil nutrients, pH, moisture and/or topography within the fields, and treating the fields according to these differences.

Precision agriculture is not a new concept, but recent interest has been propelled by advances in computer technology that make provision for the capture and analysis of spatial variability in fields, as well as advances in application technologies that allow variable-rate application of nutrients (Schnitkey, Hopkins & Tweeten, 1996). Modern technology in agriculture is an important key to success, and agriculture in general is dependent on innovations that are rapidly changing, thus presenting constant challenges to farmers. Farmers must keep up with the changes that may be beneficial to their farming operations (Roberson, 2000). Precision agriculture is used chiefly to provide site-specific data about the soil and its characteristics, which have to be processed into useful information. This information provision is facilitated by a set of composite technologies.

2.3.1 Precision agriculture technique

Precision agriculture technologies include Global Positioning Systems (GPS) receivers; Geographic Information Systems (GIS) databases; grid soil sampling; variable-rate application technologies for fertilizer, seed, lime, herbicides and pesticides; yield monitors and mapping; remote sensing imagery and proximate sensors, as well as auto-guidance systems and computer hardware and software. These techniques are interrelated, and the functionality of one component often depends on the other. A brief discussion of each of these techniques is provided in the sub-sections that follow. Variable-rate application technology, which is the focus of this study, will be discussed separately and in greater detail.

2.3.1.1 Global positioning system

The basis of precision agriculture is the provision of location-specific data, which are made possible by the GPS. This system comprises of a set of 24 satellites in the earth's orbit that sends out radio signals, which can be processed by a ground receiver to determine a specific position on earth (Rains, Thomas & Velledis, 2001). The satellites, which orbit the earth at very high altitudes, each circle the earth twice a day. A low-energy signal containing a data message is continuously transmitted by each satellite. GPS receivers everywhere – on the ground, at sea or in the air – can read and interpret these data messages. Of the 24 orbiting GPS satellites, a minimum of six to eight should be directly "visible" to a GPS receiver antenna at any point in time. Obstacles such as trees or buildings can block one or more satellites in some areas. Since satellites are not geo-stationary, some move out of view at certain times, while others come into view (Morgan, Parsons & Ess, 2000).

The GPS was originally developed as a navigational aid for military and civilian purposes. The GPS provides a horizontal position accuracy of 10 to 15 m, which implies a 95% probability that the given position will be within 10 to 15 m of the true position (Blackmore, 1994).

Differential GPS (DGPS) provides even greater accuracy, as more precise information can be obtained to sub-metre level (Trimble, 2005). DGPS receivers are adequate for tracking field positions, and this is the type of receiver found on most agricultural tractors and harvesters. To get a clear view of the sky, receivers are mounted on top of the tractor or harvester cab, or some other high point. The primary need for a DGPS system is the ability to repeatedly return to a particular point (Rains *et al.*, 2001). At harvesting time, a farmer can assess the relationship between the yield and the amount of fertilizer applied at a specific point, with the aid of coordinate points captured by DGPS. This system is required for mapping yields, field boundaries, weedy areas, or soil sampling sites, and for using variable-rate application and seeding equipment (Morgan *et al.*, 2000).

2.3.1.2 Geographic information system

Data collected by geographic positioning systems must be processed into a format useful for decision-making. This is made possible by the Geographic Information System (GIS), which is software that imports, exports and processes spatially and temporally distributed data (Rains & Thomas, 2000). *Geography* represents place, space and time. *Information* represents data and their interpretation in decision-making, while *System* represents analysis and presentation. GIS involves four sub-systems: data input, data storage, data manipulation and reporting. Maps can be computerised effectively with this system. Integration of layers of spatial information can be done with GIS through registration, and it also has the ability to uncover possible relations that would not otherwise be obvious. Registration is a process of transforming one layer of spatial information to match a second layer (Nelson *et al.* 1999). The advantage of GIS over traditional maps is the ability to combine maps to produce new maps that show interactions between factors such as yield and pH (Blackmore, 1994).

There are two forms of GIS data, namely *vector* and *raster*. A database organises and manipulates map features such as points, lines and polygons in vector data sets. Yield data points and polygons constituting management zones form vector data sets. In raster data sets, the data are organised as a matrix of numerical values and referenced spatially by row and column position. Both forms of data can be handled by most GIS software (Nelson *et al.*, 1999).

2.3.1.3 Grid sampling

Grid sampling is one of the facets of data collection required for precision agriculture. To ascertain whether a variable-rate application is needed, soil sampling is done to determine the rates of the required inputs, as well as the location where they are to be applied (Franzen, 1999). The representivity of the soil sample on which the recommendation is based, determines the accuracy of a fertilizer recommendation for the area (Johnson, 2001; Rains *et al*., 2001).

Grid sampling is a method of dividing a field into blocks of about 0.5 to 5.0 ha, and sampling soils within those grids to determine appropriate application rates. The 0.5 to 5.0 ha is a range used in the industry, but it may be larger or smaller. Several samples are taken from each grid close to the waypoint, mixed and sent to the laboratory for analysis. This soil analysis mechanism detects variability in nutrients and pH, and GPS associates latitude and longitude with this information. Maps of pH and nutrients can be produced in GIS software and a computer card can be prepared, which is then read by a fertilizer/seed applicator equipped with variable-rate application technology (Rains *et al*., 2001). The assumption behind grid sampling is the possibility of predicting values in other parts of the field on the basis of the sampled points.

In grid sampling, a systematic approach is used, and the technique assumes that there are several random patterns in a field. Franzen (1999) identifies a method to use for choosing grid sampling criteria. These criteria should be applied, for instance, in cases where the field history is unknown, fertility levels are high due to high rates of fertilizer application, there is a history of manure application, small fields have been merged into larger fields, and non-mobile nutrient levels are of primary importance.

In order to ensure that the sample accurately reflects the fertility of the area in question, certain variables must be properly considered. These include the spatial distribution of samples across the landscape, the depth of sampling, the time of the year when the samples are taken, and how often the area is sampled. The degree of variability within a given area determines the sample distribution. A systematic approach such as grid sampling is best for non-uniform sites. When soil samples are taken for nutrient recommendation, the depth should be the same as that used for the research on which the recommendations were based (Johnson, 2001). It is shown in the study of Rains *et al*. (2003) that an inaccurate soil depth significantly changes the measured properties. The recommended sampling time is after harvesting or before planting. Sampling during the growth season is not recommended, as it may give erroneous results due to the effects of crop uptake and other processes. It is advised that sampling should be done at the same time of the year and at the same spot every time a particular field is sampled, in order to facilitate tracking of trends in soil test values over time (Johnson, 2001).

2.3.1.4 Yield maps

Yield maps supply farmers and other users with visual information of the yield data collected through other precision agriculture components. Yield monitor GPS generates geo-positioned databases and site-specific yield maps. Yield maps are produced by processing data from adapted combine harvesters that are equipped with GPS integrated with yield-recording systems, with GIS as the processing mechanism (Blackmore, 1994). Yield mapping involves recording the grain flow through the combine harvester, while recording the actual location in the field at the same time (Dampney & Moore, 1999). Yield mapping makes it possible to determine spatial variation in yield within a field. Two important pieces of information, yield variability and yield production level, can be obtained immediately from yield maps.

Yield maps obtained from yield monitoring systems provide one of the most powerful sources of information for operating precision agriculture. Precise location of high- and low-yielding areas within fields can give a clear direction on where to sample in order to identify limiting factors for the maximal production of a crop (Schnug *et al.*, 1998). Yield maps are also valuable in identifying and targeting areas for investigation and treatment by precision agriculture practices, and subsequent monitoring of results. They provide a basis for the estimation of replenishment levels with regard to pH, P and K fertilizers. Information from yield maps can be applied diagnostically by helping to identify previously unnoticed problems, or may assist in determining input mix or other management strategies. The value of this information lies in its potential to increase production or reduce the use of other inputs (Lowenberg-DeBoer & Swinton, 1997). However, Godwin *et al.* (2002) note that yield maps alone do not provide a sound basis for determining a variable-rate N application strategy to optimise management in a particular season.

The importance of considering the quality of yield maps in identifying different yielding areas is stressed by Schnug *et al.* (1998). The variability in yield maps can stem from the true variability of crop yields or erroneous variability produced by the recording system itself, resulting in questionable validity of the information.

2.3.1.5 Remote sensing

Remote sensing is another set of precision agriculture tools that can provide visual information regarding the land characteristics. Remote sensors are generally categorised as aerial or satellite sensors that can indicate variations in field colour that correspond to changes in soil type, crop development, chlorophyll content, field boundaries, roads, and water. Aerial and satellite imagery can be processed to provide vegetative indices, which reflect plant health (Rains *et al.*, 2003). Remote-sensed images are regularly used in determining the Normalised Difference Vegetation Index (NDVI), which is related to the chlorophyll content and water absorption of the crop. With remote sensing, variations in crop structure can be monitored in near "real time", enabling mid-season agronomic decisions that can improve crop performance (Welsh *et al.*, 1999).

However, the economics of remote sensing in crop production have not been widely reported. Tenkorong and Lowenberg-DeBoer (2004) reviewed hundreds of remote sensing studies, and only 10 studies report on the economics associated with the use of this technology. Of the 10 studies, only seven report positive returns and provide adequate information on the budgeting methods and assumptions made to arrive at the estimates. However, the seven studies that provide information on the estimation, vary in terms of management variables and the use of remote sensing (VRT management subsequent to remote sensing), as well as remote sensing cost and analysis. As a result, no pattern was evident in the profitability estimates (Tenkorong & Lowenberg-DeBoer, 2004).

2.3.1.6 Proximate sensors

Proximate sensors can be used to measure soil (N and pH) and crop properties as the tractor passes over the field. According to Adamchuk, Dobermann and Morgan (2003) the sensing of soil variability is one of the essential steps required in precision agriculture. The Mobile Sensor Platform by Veris Technologies introduced an automatic pH sensor in 2003. With this technology, the soil sample is scooped and pressed against an electrode, a stabilization period of about 10 to 15 seconds is allowed, and the reading is then taken (Lowenberg-DeBoer, 2003).

The measurement of inherent field variability is increased with on-the-go sensors, in comparison with a one-hectare grid (Erickson, 2004). This system can be efficient in countries with high labour costs associated with grid soil sampling. At a speed of 10 to 12 km per hour in rows 18 m apart, approximately 10 readings per hectare – or one reading every 18 m – can be taken by an on-the-go pH sensor. The end result is a card that can be used in the variable-rate application of lime on a mapped field. Adamchuk *et al.* (2003) consider automated soil pH mapping a promising alternative to the conventional sampling methods to determine variability, as it increases sampling density. Erickson (2004) points out one limitation of on-the-go soil pH sensor, namely the measurement of the active pH, and not the reserve acidity. The reserve acidity, amongst other factors, determines the amount of lime that needs to be applied.

Lowenberg-DeBoer (2004) indicates that a number of real-time sensors for N application are available. These sensors take two forms: Firstly, the handheld crop sensing units, which allow measurement at different sites within a field in order to ultimately generate an application plan. This procedure still requires management time to process an application map. The second type of real-time sensors, which are mounted on application equipment, solve the management time problem. An input application is controlled by the sensor data entered into the software, instead of a map-based application system developed by a farmer (Lowenberg-DeBoer, 2004).

2.3.1.7 Auto-guidance systems

Auto-guidance systems depend entirely on GPS technology, and require a base station located on or near the farm, a rover unit for each tractor, a computer and its software. Satellite signals are sent to these systems every few seconds, and the accuracy of these signals is improved by base station correctional signals (Lewis, 2003).

Auto-guidance technology is available with two accuracy levels. Firstly is the differential GPS (DGPS) with a 10 cm accuracy; and secondly, the real-time kinetic (RTK) GPS with about an accuracy in the region of two centimetres. Auto-guidance systems allow farmers to maintain straight rows during farming operations, and to come back to the same rows in the next season. These systems make more precise input applications possible. Watson and Lowenberg-DeBoer (2002) identify, amongst others, the following benefits associated with auto-guidance technology: fewer skips and overlaps, with more precise planting, fertilizing and tilling; less compaction, which leads to higher yields; reduced operator stress and fatigue, which improves efficiency and extends the potential hours of operation. According to Watson and Lowenberg-DeBoer (2002), larger farms are more likely to benefit from auto-guidance systems, and these systems allow farm size expansion with the same set of equipment**.**

2.3.1.8Computer hardware and software

Computer support is required in order to analyse the data collected by the other precision agriculture technology components and to make it available in usable formats such as maps, graphs, charts or reports (Manitoba Agriculture & Food, 2001).

The combination of precision agriculture technologies discussed in previous sections, provides valuable information to the agricultural sector regarding the components and mechanisms of this technology.

2.3.2 Precision agriculture information

Information is probably the most valuable resource for a modern farmer, and knowledge of precision agriculture information may even be the most important tool at his disposal. In all phases of production, from planning through to the post-harvest phase, timely and accurate information is essential (Roberson, 2000). Recent technical innovations for precision agriculture provide farmers with spatially referenced data on the nutrient content and soil quality of fields (Isik *et al*., 1999). Information available to farmers in this regard includes crop characteristics, soil properties, fertility requirements, weed populations, insect populations, plant growth response, harvest data, and post-harvest processing data (Roberson, 2000). Soil properties and fertility requirements are more relevant for this study, as they constitute variables considered in explaining yield variation.

Precision agriculture is also described by Lowenberg-DeBoer and Swinton (1997) as information technology applied to agriculture. Collecting and analysing information costs money – even more so in the case of precision agriculture information. Although precision agriculture information is not as easily quantifiable as other inputs in the production process such as seed, fertilizer and pesticide, its economic role can be analysed in much the same way as that of conventional inputs. Information and management are related. Information can be a significant source of strategic competitive advantage, resulting in better performance and higher profits compared to other farmers (Lowenberg-DeBoer & Boehlje, 1996). Farmers who possess detailed and precise information, collected through precision agriculture activities, have a competitive edge over farmers who employ conventional farming methods and have only limited information at their disposal about their fields and the variability that exists in them.

The agricultural role players in South Africa have also realised the importance of precise information in the success of the country's agriculture. An endeavour has been made to develop an Agricultural Geo-referenced Information System (AGIS). This is collaboration between the National Department of Agriculture, the Agricultural Research Council and the Provincial Departments of Agriculture. The vision is to acquire, coordinate, describe and manage all relevant information on agriculture in South Africa in an integrated, geo-referenced information system in order to provide information at national, provincial and local level. Some of the decision support systems provided, which are accessible through the Internet, provide decision-makers with timely, current and relevant information to secure a competitive advantage in international markets through more effective decision-making (Rust, 2005). Currently, AGIS gives farmers access to information on the location of towns, rivers, roads, administrative areas and farm boundaries, as well as information on soils, natural vegetation, climate and land capability on a national scale. For instance, the foot and mouth outbreak map indicates the areas affected by recent foot and mouth disease outbreaks.

Precision agriculture constitutes the information and diagnosis component of an entire farm management system. Yield monitoring and other site-specific information provide the farmer with an improved ability to diagnose crop production problems (Manitoba Agriculture & Food, 2001). The value of this information lies in its potential to bring about increased production or reduced input use. The economic value of gathering and analysing information is demonstrated in the study of Lowenberg-DeBoer and Swinton (1997), where the profitability of precision agriculture is analysed. Malzer *et al*. (1999) states that the benefits gained from precision agriculture are associated with how the manager or farmer integrates the information into a package that interprets the field variability into different management strategies.

Precision agriculture information can take different forms: data generated at the end of the production season, such as yield output data and crop input needs data, as well as data generated during the growing season with regard to nutrient deficiencies, insect and weed infestation, and the general growth of the crops. A number of factors can influence the yield of crops in a given location. Although little control can be exercised over factors such as soil texture, climate and topography, they must be considered in view of their effect on yield. Other factors such as soil structure, pH levels, macro and micro nutrients, available water, water-logging, weed competition, pests and diseases can be manipulated in a spatially variable manner, and can give rise to economic benefits (Godwin *et al*., 2002). Consequently, careful analysis of yield data is essential, often over a period of years, before inferences can be made about appropriate management actions. In contrast, data on crop nutrient needs usually result in a clear prescription for action. Even though the prescription may be clear, the potential economic value thereof requires analysis in order to inform farmers on the bottom-line profitability of the investment (Lowenberg-DeBoer & Swinton, 1997). Thorough analysis of the information is as important as the information itself with a view to determining bottom-line profitability, which is the main interest of farmers and agri-business at large.

2.3.3 Profitability of precision agriculture

The profitability of precision agriculture tools is an important consideration for farmers and agri-business with a view to determining whether they are heading for a promising future or a technological dead end (Lowenberg-DeBoer & Swinton, 1997). The profitability of precision agriculture is the single most important consideration with regard to this technology, and determines whether it will be adopted or not.

Lowenberg-DeBoer and Swinton (1997) summarised the results of 17 field crop precision agriculture profitability analyses. Overall, five studies found that precision agriculture is not profitable, six produced mixed or inconclusive results, and the other six indicated potential profitability. Profitability as used in this context implies that switching to precision agriculture results in higher net returns than whole-field management. However, direct comparison between different studies is not possible as different assumptions were made in different studies, particularly with regard to cost accounting (Lowenberg-DeBoer & Swinton, 1997).

Lambert and Lowenberg-DeBoer (2000) also reviewed and summarised 133 publicly available studies on the profitability of precision agriculture. Sources of the reviewed articles included scientific journals or proceedings (86%) and non-technical or non-refereed magazines and monographs specialising in agri-business services (14%). Of the reviewed articles, a total of 108 reported on the economics associated with the technology, with 63% indicating positive net returns and 11% negative returns for a specific technology, while the remainder reported inconclusive results. Calculations carried out to determine economic returns in these studies varied, particularly with regard to costs and returns included.

A distinction was made between articles authored by economists (63%) and non-economists (37%). Seventy-three percent of the articles authored by economists indicated positive benefits from precision agriculture, 16% indicated mixed results, and 11% negative results. With regard to the articles authored by non-economists, 66% reported positive results, while mixed results were reported in 22% of the articles.

Yield estimation techniques played a role in the analyses. A slightly higher percentage of positive results was obtained in studies using response functions (60%) or simulation (75%) to estimate yield, relative to studies using field trial data (67%). The popularity of reviews varied between different crop categories, with 54 of the articles focusing on economic returns generated by precision agriculture technology in maize studies, followed by wheat (13%), sugar beet (3%), potatoes (4%) and soybean (3%). VRT was the most commonly used precision agriculture component, and was used in 73% of the studies reviewed. Nitrogen management constituted 21% of VRT, followed by VRT-P at 5% and

VRT-pH at 3%. Variable-rate seeding and irrigation followed VRT fertilizer management at 7% and 2%, respectively.

Variables considered in the partial budgeting technique included time scale and discount rate, input and VRT/PA costs, and human capital and information costs. Farm inputs in the budget analyses were included in 99% of the reviewed articles, while 81% included precision agriculture technologies. Human capital costs were included in 31% of the articles, while information costs were included in 44% of the reports (Lambert & Lowenberg-DeBoer, 2000).

The precision agriculture review of Lambert and Lowenberg-DeBoer (2000) was updated by Griffin *et al.* (2004b). About 243 articles were reviewed, and 210 of these reported benefits or losses. Of these 210 reviews, 68% reported benefits from the given precision agriculture technology, and half of these were authored or co-authored by economists. GPS as a stand-alone featured in 6.4% of the articles, VRT with GPS in 4%, and with yield monitor and seed in 3% of the articles.

Economic returns on maize alone featured in 37% of the reviewed articles, with 73% of the articles reporting some benefits associated with PA. Maize frequency was followed by wheat (11%), with benefits being reported in half of these cases. Maize and soybeans were also mentioned in 9% of the articles, with benefits being reported three-fourths of the time. Benefits were reported in all soybean, barley and oats studies, while no benefits were reported in maize and cotton combinations.

As in Lambert and Lowenberg-DeBoer (2000), partial budgeting components reported differed. Farm inputs were included in the budget analyses in 71% of the articles, while 62% included precision agriculture costs. A quarter of the articles mentioned equipment costs, and 40% mentioned yield monitors. Environmental benefits were reported in one third of the articles. With regard to capital and information costs, approximately 21% of the reviewed articles included capital costs, while information appeared in 34% of the reviewed articles (Griffin *et al*., 2004).

It should be borne in mind that there are external factors besides the soil-plant relationships that have an impact on the profitability of precision agriculture. According to Manitoba Agriculture and Food (2001), crop prices affect the use of variable-rate application technology. VRT may be more appropriate for high-value crops such as beans and potatoes, as opposed to bulk cereals and oilseed crops.

Moss and Schmitz (1999) also support the notion that the value of precision agriculture is related to the price of the product. According to these authors, the value of precision agriculture can increase with the price of maize, assuming that N price remains constant. This arises from the fact that, as the value of higher yields increases, the allocation of the marginal unit of fertilizer between land types becomes more important. On the other hand, the value of precision agriculture declines as the price of fertilizer increases together with the price of maize. An increase in profitability should be expected as managers, consultants and researchers gain experience, and as the amount of data collected increases. As more precision agriculture practices are adopted, the costs of precision agriculture are expected to decrease due to economies of scale and/or a reduction in costs (Shatar, 1998). According to Schilfgaarde (1999), the expectation is that precision agriculture will increase crop yields and enhance net returns from farming, while reducing environmental damage at the same time.

Lowenberg-DeBoer and Swinton (1997) conclude that, since yield and input use changes will vary from one farm to the next, it is difficult to make a general statement about the profitability or non-profitability of precision agriculture. The profitability of any given precision agriculture technology and the factors involved may be site-specific, and what works in one area may not necessarily work in another. As a result, precision agriculture should be evaluated on a farm-by-farm basis (Lowenberg-DeBoer & Swinton, 1997; Malzer *et al.*, 1997; Shatar, 1998). The view of Anselin, Bongiovanni and Lowenberg-DeBoer (2004) is that the profitability of VRT crucially depends on the model specification used, as all the spatial models they investigated consistently suggested profitability, while non-spatial models did not.

An entire section is devoted to the evaluation of variable-rate application techniques later on in this chapter. Profitability analysis of variable-rate application technologies for different inputs can facilitate in-depth analysis of precision agriculture, and provide insight on any risks that may arise.

2.3.4 Risk in precision agriculture

Many new technologies have risks associated with them, and precision agriculture is no exception. It is important that these risks be identified timeously in the decision-making process. Although precision agriculture provides some benefits, these benefits may go hand in hand with increased risk. Making the wrong decision for a field can result in substantial costs. These risks can be minimised by understanding some of the factors that might influence nutrient utilisation by crops in order to make the right decision for a particular situation.

Lowenberg-DeBoer (1999a) discusses some of the sources of risk associated with precision agriculture. Yield and returns are potentially increased by precision agriculture, but the possibility of crop failure is also present. In a bad crop season, payments for soil sampling, VRT application of inputs and other services may increase losses. Increased risk also arises from human and technological factors. Someone on the farm must have the necessary skills to operate the equipment and interpret the collected data. If that person is no longer available, someone else must be trained, and it will take time for this substitute to develop the skills and the necessary experience to master the technology and its performance. Interpretation of data often requires site-specific knowledge that grows with experience. Human risk also involves the layering of maps, which opens up the possibility of making wrong decisions.

Lowenberg-DeBoer and Swinton (1997) view risks in precision agriculture as arising mainly from changes in yield variability, and from business risk related to obsolescence. Technological risks come mainly in the form of obsolescence, as the technology changes so rapidly (Lowenberg-DeBoer, 1999a). As precision agriculture technologies are still in their infancy, the resulting technological obsolescence can make it difficult for farmers to recover their sunk cost if the investment is to be liquidated due to a downward turn in revenues (Isik *et al.*, 1999). The possibility that equipment and software will not be supported in the future is one of the most important risks related to obsolescence. It is therefore important to consider this factor when choosing suppliers and drawing up long-term plans (Lowenberg-DeBoer & Swinton, 1997).

Farm planning can be both easier and more complex with precision agriculture. There is an abundance of map data that can be utilised in management decisions, requiring more effort and work in interpreting the data, which increases the risk of misinterpretation (Goddard, 1997). If it is found that in the analysis of precision agriculture that the annual gross margin outweighs the annual capital costs (with or without adjustment for environmental benefits), and the expected profitability of the proposed precision agriculture technology has been confirmed but projections on risk and feasibility are still outstanding, more analysis is essential (Malzer *et al*., 1999).

It has been hypothesised by Lowenberg-DeBoer (1999a) that precision agriculture technologies such as GPS, GIS and VRT may be helpful in managing risks. This hypothesis is based on the concept that, with precision agriculture technologies, more and better information is available and control over crop cultivation conditions is increased, thereby reducing variability in net income. Early yield estimates can assist producers in marketing by providing better information on the growing crop and increasing confidence in contracting or hedging (Manitoba Agriculture & Food, 2001; Lowenberg-DeBoer, 1999a), thus reducing price risk.

Precision agriculture can also assist in measuring relative yield risk by comparing yield variance to a standard level of single-rate application. Lowenberg-DeBoer and Swinton (1997) state that there is no evidence that precision agriculture increases yield risk. If precision agriculture can reduce the spatial variability of yields, it may even be considered a risk-reducing technology. The numerous other benefits associated with this technology will be discussed in the next section.

2.3.5 Benefits of precision agriculture

Before any decisions are made about the adoption of this technology, it is essential that an extensive analysis is done regarding the feasible benefits. Although the farmer’s interest is the bottom line, pursuing an endeavour becomes more lucrative if there are other spill-over benefits associated with the technology. Three main ways in which precision agriculture can increase financial returns are identified by Shatar (1998). These are reduction in fertilizer, chemical and seed costs, increase in yields and improvement in the quality of the crop. Precision agriculture allows improved economic analysis. A computerised method of capturing, storing and analysing field records provides several benefits, as detailed analysis of the farm production and management activities can be carried out (Goddard, 1997). By recording field operations and input applications, precision agriculture technology makes it possible to track the production process to help guarantee food safety, and to document how food was produced, thereby facilitating identity tracking and identity preservation. Compliance with environmental regulations can also be documented with precision agriculture (Harris, 1997). With GPS guidance, skipping and overlapping with regard to input applications are reduced, making longer working hours and higher labour productivity possible. GPS guidance also allows accuracy at higher speed, is less affected by weather and reduces operator fatigue and eye strain, amongst other benefits (Lowenberg-DeBoer, 1999b; Lewis, 2003).

According to Godwin *et al.* (2002), precision agriculture can facilitate the identification of problems that could result in significant crop yield penalties, such as fertilizer application errors and water-logging. Lost revenue can be calculated, and the resulting impact on cost benefit can be determined. This provides a basis for making informed management decisions. Godwin *et al.* (2002) stress the importance of rectifying these problems before implementing the spatial application of fertilizer and other inputs. Lowenberg-DeBoer and Boehlje (1996) reaffirm the essence of precision agriculture, which is acquiring additional data on production processes and converting such data into useful information that can be used to manage and control these production processes.

Precision agriculture also presents some environmental benefits. These benefits can be considered if the annual gross margin from precision agriculture cannot cover annual capital costs, and can be reaped in two ways. Firstly, some environmental benefits lead to direct financial gains. Secondly, there are other environmental benefits that hold social or psychological value to certain managers, and not monetary value (Isik *et al.*, 1999). If one or both of these benefits appeal to a particular manager or farmer, adoption of the technology can be enhanced.

The reduction in excess N application attributable to precision agriculture can result in certain public environmental benefits without any loss in yields to the farmer (Babcock & Pautsch, 1998). In the simulation model of English, Mahajanashetti and Roberts (1999), which estimates maize yield and N loss response functions for applied N, considerable environmental benefits are also realised as N loss from leaching, sub-surface flow and surface runoff is reduced. Bongiovanni and Lowenberg-DeBoer (2004) analysed the effect of restricting N application for environmental stewardship and complying with environmental legislation. The N applications are constrained at different levels. With the N application constraint of 60 kg/ha, VRT application reduces profitability by $0.15/ha while profitability is reduced by $0.45/ha when using the uniform application. The conclusion is that with VRT application of inputs, the profitability may be maintained while fewer inputs are used and environmental damage is limited. VRT N application maintains profitability even in a scenario when N is restricted to less than half of the profit maximising uniform rate (Bongiovanni & Lowenberg-DeBoer, 2004).

Bongiovanni and Lowenberg-DeBoer (2004) also reviewed several studies on the economic and environmental impacts of the variable-rate N application in crop production. Of the 15 reviewed studies, only two did not show any environmental benefits from variable-rate application of N. Although environmental benefits make precision agriculture a more attractive package, the chance to increase profits will be what drives most farmers to adopt this technology (Isik *et al.*, 1999). Similarly, the external effects arising from environmental benefits are not felt by individual farmers, making these benefits of little significance. Environmental benefits are of little or no value to farmers in developing countries such as South Africa because the environmental legislation is not yet well-developed and consumers are less concerned about the environment.

The ability to use the spatial variability of soils to their advantage is another motive for farmers to adopt the variable-rate application technologies of precision farming. Spatially detailed soil and crop information can reveal, with reasonable accuracy, where increased fertilizer application would be profitable and where it would not be productive. According to Weiss (1996), obtaining and applying spatially refined information entails both costs and benefits. It is therefore important to ascertain under what circumstances enhanced profitability will be the net result, and what environmental benefits will result from better-targeted chemical use.

Weiss (1996) and Schnitkey *et al.* (1996) assert that precision farming has economic or environmental benefits only in the presence of spatial variability. English *et al.* (1999) support this notion by concluding that the magnitude of VRT lucrativeness and N loss reduction is influenced by spatial variability, and that adoption of VRT is not favourable where yield variability is low. Moss and Schmitz (1999), as well as Isik *et al.* (1999), reiterate the importance of soil variability. They assert that the gross benefits gained from precision agriculture will be enhanced by greater variability in the distribution of the primary productivity of the soil. When a small percentage of low-quality soil is present, VRT application of fertilizer on this soil is of little benefit. In the same way, the benefit of applying more fertilizer on high-quality soil is small when there is only a small percentage of high-quality soil. It is therefore important for agronomists and agricultural economists to establish what spatial configurations of soil and crop characteristics display enough variability to make precision farming profitable, namely to increase crop yield, improve crop quality or reduce fertilization to such an extent that it will more than compensate for the increased cost of obtaining, interpreting and applying spatial information (Schnitkey *et al.*, 1996).

The value of precision agriculture depends on the crop response to the variation at hand. For instance, the soil fertility variation that occurs mostly in the high fertility range may result in little crop response, and hence limited economic benefits. On the other hand, profitability is possible under conditions of slight variability. Variable-rate seeding is potentially profitable when the low-yield proportion (<6 tons/ha) is small (Lowenberg-DeBoer, 1999a). Bullock *et al.* (1998) also found that, in view of the knowledge available about the response to seeding rate in every section of the field, variable-rate seeding can be profitable under conditions of limited variability if the equipment and service costs do not exceed $9.83/ha.

2.4 VARIABLE-RATE TECHNOLOGY

Variable-rate technology, the aspect of precision agriculture that is the focus of this study, entails the precision application of inputs. Inputs are varied throughout the field according to pre-determined yield potentials or other guidelines. VRT includes computer controllers that make provision for variation in inputs such as seed, fertilizer, lime, herbicides and pesticides. Application rates are varied as areas of different potentials or problems, which warrant different rates, are encountered (Tiffany & Eidman, 1998). Applications can be controlled by the farmer, or by a geo-referenced card prepared in GPS and GIS systems**.**

Rains and Thomas (2000) examine how VRT can demonstrate the interaction of the components of precision farming, and how they fit together. The GPS tells the operator where on the field the tractor is located. The GPS also links with GIS to give the controller information about the field characteristics at that location. The yield potential or planned yield, together with the soil nutrient status, precisely indicates fertilizer application at that spot. The correct amount of fertilizer to be applied is then manipulated by the VRT applicator or controller. The effects of variable-rate application technology can be observed and evaluated on the basis of the yield data collected by other precision agriculture mechanisms, mainly through the yield monitor. Before such collected data can be analysed, they should be free of any erroneous matter that could bias the results. This is achieved through data cleaning.

2.4.1 Cleaning of VRT data

Monitoring the yield can provide valuable information concerning the extent and location of variability in a field, but this is only possible with clean data, free of any errors. Abnormally high and low yield values can be recorded during data collection, and the removal of this erroneous data is imperative. Erroneous data can result from rapid speed change or not cutting the full width of the header, as well as from a yield sensor that is incorrectly calibrated (Drummond, 2005; Kleinjan *et al.*, 2004). False high readings can occur when a combine stops moving and the crop that is being harvested, continues to be threshed. Since yield is calculated as the crop quantity in tons per distance travelled, a false high-yield reading is obtained because the distance travelled is very small when the combine has stopped or almost stopped, but the crop continues to be recorded at a regular rate. Speed variation can also lead to false yield readings.

When yield data is used in a decision support system, any data that can be considered flawed, must be removed. Removal of incorrect data can have a significant impact on the analysis of data by allowing valuable comparison of yield data with other information layers within the decision support system. Correct data can allow layering between yield, pH, soil type and inputs applied in the production process, as well as investigation into relationships that exist.

Kleinjan *et al.* (2004) describe the actual cleaning of data as a multi-level process consisting of various phases. The first phase involves discovering errors that are easy to identify, such as the conditions that arise when the combine header is up. When the header is up and the crop is entering the combine, but the yield data are being recorded, the resulting values are invalid. In the second phase of data cleaning, yield monitor data recorded at rapidly changing speeds are identified. The above-mentioned authors use the decision rule of 15% velocity change, which means that, when there is a velocity change of more than 15% of the initial velocity from one reading to the next, the reading is excluded.

The third phase of data cleaning concentrates on the speed of the machinery. Any yield recorded at a speed of less than 1.6 km per hour is discarded. If V_i is <1.0 km per hour, then the line of data containing V_i is discarded. The logic behind this is that, when the velocity of the combine is close to or equal to 0 km per hour, the area harvested moves towards 0. Since a combine can stop moving while still recording yield, any mass recorded is divided by an area to the value of 0, resulting in infinite yield. In addition, these data are excluded because normal calibration of most modern combines and their yield monitors occur at velocities of 4.8 to 8.0 km per hour. Therefore data recorded at velocities below 1.6 km per hour are regarded as abnormal and unreliable.

The flow of grain past the yield monitor is taken into account in the fourth phase. When calibrating yield monitors, the upper and lower flow limits are set on the basis of the yield potential, historical yields and fertilization target yield. If data are recorded at flow rates outside the range that was calibrated for, the collected flow-rate data are regarded as implausible, and are therefore not used. The fifth and last phase of data cleaning involves calculating the average and standard deviations of the yield data within a particular distance constraint. Within a block (three header widths of a combine by three header widths), homogeneous data are expected. Yield data within a block that is three standard deviations from the block mean, are discarded (Kleinjan *et al.*, 2004). This phase of data cleaning can be used to clear any suspect data in grid data analysis. In the process of analysing the variable-rate application technology, data cleaning must be applied according to a correctly specified model to obtain meaningful results.

Yield Editor, software developed by Drummond (2005), performs editing, filtering and cleaning of data in a matter of a few minutes. With Yield Editor, erroneous observations can be identified and removed by looking at harvester velocity (minimum and maximum), velocity change, start- and end-pass delay, maximum and minimum yield, flow delay and standard deviation. Data from AgLeader Advanced and Greenstar text formats can easily be imported into Yield Editor, and the processed data are then exported in ASCII format (Drummond, 2005). Griffin, Brown and Lowenberg-DeBoer (2005a) provide a protocol for managing the analysis of site-specific production data with the aim of analysing field-scale experiments. Amongst other software used in the protocol, Griffin *et al.* (2005a) share their experience in using Yield Editor for filtering erroneous data. Yield Editor and its manual are available for download at http://wwww.fse.missouri.edu/ars/decision_aids.htm. Yield Editor was used for cleaning data in this study.

2.4.2 Model specifications for VRT

In order to get meaningful results from the analysis of variable-rate application technologies, possible model specifications must be scrutinised and the appropriate model must be selected. Babcock and Pautsch (1998) identify two key issues that are involved in developing a model of production decision under variable-rate fertilizer technology. The first key issue is the selection of a functional relationship between yields and fertilizer levels for a given crop, as different estimates of the value of VRT arise from different functional forms.

The second important issue in the development of the model is the selection of a field attribute that can be used to guide fertilizer rates. An ideal VRT fertilizer management would ensure optimal fertilizer rates in a field by taking all factors that affect optimality into consideration. These factors include cropping history/historical yields, manure application, previous fertilizer practices and inherent soil characteristics. To make VRT fertilizer application possible, variations in one or more of these attributes should be measurable, and should be used to adjust application rates across a field. An attribute that is easy to measure and will reliably predict intra-field fertilizer rates, is ideal (Babcock & Pautsch, 1998). As English *et al.* (1999) point out, the main principle guiding a profit-maximising farmer is the expected economic benefits of implementing VRT, and various models can be used to estimate this, as detailed below.

A single production function cannot represent agricultural production under all environmental conditions. The functional form and the magnitudes of its coefficients vary according to soil, climate, type and crop variety, different resources and the magnitude of other fixed inputs. As a result, a specific type of equation that expresses production phenomenon automatically imposes restrictions or assumptions with regard to the relationship involved (Heady & Dillon, 1961).

2.4.2.1 Linear Response and Plateau

Linear Response and Plateau (LRP) models exhibit a proportional response with a constant marginal product up to a point where further input application does not result in yield increase, but rather a yield plateau (Mortensen & Beattie, 2005). LRP function can be used to determine the economically optimal fertilizer rates. Below the plateau level of production, a linear response and plateau production function has a constant marginal product, while a constant level of total physical product is obtained after reaching the plateau. Under this production function, the optimal decision would entail applying N at a rate that is adequate to achieve the plateau. This would be economical if the marginal physical product (MPP) of N along the plateau path is greater than the ratio of the price of N to the price of maize. If the latter is greater, it would not be optimal to apply N at a rate adequate to reach the plateau (Moss & Shmitz, 1999). There is great fluctuation in the price of maize, and at the current (2006) low prices, care should be exercised in using this functional form to determine optimal fertilizer rates.

A LRP relationship between yield responses to applied N fertilizer at field level is assumed in the study of Babcock and Pautsch (1998). A linear relationship between expected yield and applied N is obtained through integration over all sub-field units. Yield potential in a field, as deduced from the management history, is used as a key attribute to guide fertilizer rates. Yield potential is related to traditional soil maps reflecting field traits such as slope, hills, valleys, and clay and sand content. The assumption is made that yield potential and optimal fertilizer rates are predicted accurately on the basis of soil maps, and that yield potential predicts optimal fertilizer rates. However, Babcock and Pautsch (1998) admit that this is not always the case, and that inaccurate predictions can be made, resulting in an overestimation of the value of implementing VRT.

Even though the change in profitability associated with switching over from single-rate technology to VRT is determined in the study of Babcock and Pautsch (1998), a linear function is not very appropriate. Since the MPP of the input is constant, the production processes are, in fact, inappropriately represented due to the fact that diminishing returns – a characteristic of crop production functions – are not allowed.

Small changes in independent variables for any fundamental function can be estimated reasonably well by linear functions (Lau, 1986). However, they do not work well for other purposes. An example pointed out is a production function where perfect substitution among different inputs and constant marginal products is implied. The phenomenon of diminishing marginal returns cannot be well represented by linear production functions. In addition, linear production functions have the perfect substitution property, with the implication that, in virtually all cases, only a single input will be employed and a minor change in the relative prices of inputs will result. This will cause a shift from one input to another – an inference that is inappropriate for agricultural production, which requires more than one input and has no perfect substitution between inputs. This functional form would therefore not correctly fit the kind of data for this study, namely typical crop production data.

2.4.2.2 Quadratic Plus Plateau

The Quadratic Plus Plateau (QPP) model exhibits a diminishing MPP up to a point of maximum yield, becoming horizontal thereafter (reaching a yield plateau). The maximum attainable yield from the application of an input (N) occurs over an extended plateau, beginning at N^{max} rather than at a single point, as in quadratic polynomial response functions (Mortensen & Beattie, 2005). The QPP function can describe the response of yield to fertilizer nutrient uptake. The empirical study of Voortman and Brouwer (2001), wherein the spatial variability of millet yield is modelled, shows that a quadratic formulation conforms best to the data. It describes the relationship between millet yield and N, P and K most effectively. Furthermore, the satisfactory explanatory power of 81%, which explains yield variation, is achieved.

The approach followed in a study by English *et al.* (1999), which is aimed at examining the economic feasibility of VRT nutrient application in relation to changing spatial variability and weather conditions, is the same as that used by Babcock and Pautsch (1998), except that the quadratic instead of the linear response function is specified. The net return differential between VRT and uniform rate technology (URT), which is basically the return to VRT, is estimated for each soil type in the field, and a summation is made for the entire field. Taking custom charges (C) into account, VRT produces the expected economic gains if return to VRT is greater than C (English *et al.*, 1999).

In the study by Alivelu *et al.* (2003), the QPP model fits the response data with less systematic bias than other tested models. The QPP function was also used in the study of Bongiovanni and Lowenberg-DeBoer (1999) to estimate the response of maize and soybean to pH (English *et al.*, 1999).

2.4.2.3 Quadratic Polynomial Response

Quadratic Polynomial Response (QPR) models show a smooth increasing MPP curve followed by a unique yield maximum, which is then followed by decreasing yield. The MPP first increases at an increasing rate, and then increases at a decreasing rate. Substitution among the inputs is an assumption of QPR functions (Mortensen & Beattie, 2005). Like other functional forms, a QPR is used to describe crop response to applied fertilizer.

Malzer *et al.* (1999) established field experiments to evaluate the potential economic benefit associated with site-specific N rate management. A QPR model is used to describe crop response to applied fertilizer at each level of analysis. Polynomial regression equations are calculated for the whole field, as well as for each of the sub-blocks that make up the whole-field data set. For all coefficients of the quadratic model, a wide range of values are observed. There is a strong inverse correlation (R = -0.94 to -0.97) between the linear coefficients and the quadratic coefficients. It is concluded that most of the response variability is explained by the intercept (grain yield at 0 N) and the slope of the linear coefficients, in view of the correlation that exists in this regard (Malzer *et al.*, 1999).

Lambert, Lowenberg-DeBoer and Bongiovanni (2004) use, amongst others, a QPR models in their evaluation of statistical methods for precision agriculture data. Maize N response data from another study were used. Statistically significant coefficients for N response by topography were found, confirming the presence of spatial variation in N response by landscape zones.

Polynomial models have a simple functional form, and are computationally easy to use. However, they have poor extrapolatory and interpolatory properties. Good fits can be obtained within a range of data sets under consideration, but deteriorate rapidly outside the data range (NIST/SEMATECH, 2004). The high inter-correlation among the independent variables in quadratic functions may result in difficulty in estimating the betas (*b*'s). This requires the implementation of tests and corrections for this correlation in order to produce unbiased estimates.

Where Bongiovanni and Lowenberg-DeBoer (1999) use the quadratic model to estimate the maize response to N, the estimated coefficients have the expected signs, leading to reasonable estimates of maximum physical yields. Quadratic models are used in many other studies, including Lambert, Lowenberg-DeBoer and Bongiovanni (2002, 2003 & 2004), as well as Hurley *et al*. (2003), to estimate maize response to N. Since the quadratic function conforms best to the data in the studies that relate well with this study, it is used as a Sensitivity Test Model.

Mortensen and Beattie (2005) contend that the choice of the functional form has significant practical implications for farm profitability and environmental quality. Different maximum N-norms were generated from various functional forms, and the QPR resulted in a figure that is in line with practical farming experience.

2.4.3 Evaluation of variable-rate technology

Even though current developments in application technologies allow variable-rate application of all inputs, much of the interest has focused on fertilizer application, perhaps due to the knowledge available on fertilizer-soil nutrient-yield relationships. The relative importance of fertilizer among other crop production expenses adds to this interest in variable-rate fertilizer application (Schnitkey *et al.*, 1996).

Fertilizer application is one of the important inputs in crop production, and the expenses associated with it can affect the profitability of the enterprise. Variable-rate fertilizer application aims at increasing the efficiency of fertilizer use and/or reducing fertilizer costs. According to Malzer *et al.* (1999), site-specific management (VR application) of fertilizer nutrients has the potential to substantially improve the economic return to producers by 10 to 20%. However, the potential benefits may also be associated with increased risks. This section reviews studies conducted on the analysis of variable-rate fertilizer application.

2.4.3.1 Variable-rate application of lime

Parkhomenko *et al.* (1999) conducted a study to evaluate the profitability of variable-rate fertilization with lime and potash, as well as to investigate whether precision agriculture strategies result in higher net returns compared to single-rate application, and different profitability results were obtained. The strategies under investigation include the *Ag strategy*, *Ec strategy* and *Inf strategy*. The Ag strategy, the common strategy for beginners, is VR fertilization that uses information from intensive soil sampling to establish fertilizer recommendation rates for individual grids, based on conventional agronomic recommendation rules. The Ec strategy is VR fertilization based on economic rules. It is similar to the Ag strategy, but uses economic rules as a basis for a recommended fertilizer rate for each individual grid. This follows the rule that a profit-maximising farmer tries to optimise the use of inputs in crop production by using them to the point at which the additional returns are equal to the additional costs. The economic rule is the approach recommended in the production economics literature. In the Inf strategy, a single rate is determined using information with regard to intensive soil sampling, and agronomic recommendations are used.

The three strategies are evaluated relative to Whole-field Management (WFM), which uses information regarding average nutrient levels of the field obtained from composite soil samples to determine a single-rate fertilizer recommendation rate for a whole field. In the case of lime, the highest Net Present Value (NPV) for the experimental field is achieved by the Ec strategy, followed by the Ag strategy. Both the Ec and Ag strategies result in higher NPVs compared to the WFM strategy (1.4% and 0.8%, respectively), due to the significant response of crops to proper liming. The resulting increase in yield may cover the additional cost incurred for precision agriculture. Adequate profits can not be generated from the intensive data collection in the Inf strategy to cover additional costs, and the NPV is 0.9% below WFM (Parkhomenko *et al.,* 1999).

The profitability of variable-rate application of lime as a stand-alone is also evaluated by Bongiovanni and Lowenberg-DeBoer (1999). The results are consistent with those obtained by Parkhomenko *et al.* (1999), namely that the Ec strategy is the most profitable option, giving a 4.82% improvement in return above variable cost, or $7.91 per acre per year more. The second-most profitable option in this study is the Ag strategy, with a 1.78% improvement in return above variable cost or $2.93 per acre per year more. The Inf strategy is least profitable – less profitable than the WFM strategy – because of the high rate of lime required, bringing the lowest grid cell up to the plateau yield level (Bongiovanni & Lowenberg-DeBoer, 1999).

2.4.3.2 Variable-rate application of potassium

In the same study by Parkhomenko *et al.* (1999), the highest NPV for potash fertilization is obtained with the WFM strategy. Lower NPVs of -1.4% and -2.8% are generated respectively for the Ec and Inf strategies. No economic benefits are obtained from Ec and Inf strategies relative to WFM strategy, because crop response to the VR-applied fertilizer cannot recover precision agriculture costs. The lowest NPV, 7.6% lower than WFM, is obtained with the Ag strategy. The reason is an initial high nutrient level, resulting in no agronomic need for potash application.

Schnitkey *et al.* (1996) conducted a similar study on P and K, where their “average” and “information” strategies are similar to the Ag and Inf strategies of Parkhomenko *et al.* (1999). Schnitkey *et al.* (1996) found in their study that the average return from the precision and information strategies is $3.28 per acre per year on eighteen grid-sampled fields. However, there is a significant variability in returns, ranging from a low of $0.74 per acre to a high of $5.36 per acre. The average difference between the information (Inf) and the average (Ag) scenarios is $5.74. Schnitkey *et al.* (1996) conclude that both aspects of the technology (information gathering and precision application) have economic benefits. Precision fertilizer application does not have positive benefits in all the fields, as it generated low returns in some fields (Schnitkey *et al.*, 1996).

2.4.3.3 Variable-rate application of nitrogen

Babcock and Pautsch (1998) developed a fertilizer decision model to estimate the potential value of moving from uniform to variable N rates in Iowa corn production. Babcock and Pautsch (1998) conclude that, if some soil types are undersupplied and others oversupplied, the value of VRT lies in yield increases as well as input cost saving. In their model, Babcock and Pautsch (1998) illustrate that, if the optimal WFM application is used and optimal fertilizer rates and soil types are linearly related, there will be an oversupply on 66% of the acres, a 4% undersupply, and correct fertilization on 30% of the acreage. The optimal single fertilization rate will equal the optimal VRT application for the entire field only if the field has one soil type. All fields exhibit some type of variability in that study, and the optimal single rate is equal to the VRT rate on 30% of the soil; no change was experienced due to the implementation of VRT.

It is assumed in the study of Babcock and Pautsch (1998) that physically optimal N rates are relatively responsive to maximum yields. The results indicate that switching to VRT will increase gross returns over fertilizer costs by $4.44 per acre over the entire study area. The vast majority of the increase (86%) in gross returns when switching to VRT arises from the reduction of the excess fertilizer applications. Another source of increasing profit with VRT arises from reducing the under-application of the N fertilizer. Applying more N fertilizer where it is needed will increase yield and farm profit. Only 14% of the increase in profits is attributable to increasing yields in the study area. The implication is that a large amount of land is oversupplied with N fertilizer when using SR application strategy.

In view of Godwin *et al.* (2002), variable-rate application of fertilizer, particularly N, can improve the efficiency of cereal production by managing variations in the crop canopy. In their research, Godwin *et al.* (2002) discovered that seven out of eight treatment zones offer positive economic returns to variable-rate N application, with an average benefit of £22 per hectare. According to Godwin *et al.* (2002), the benefits of VRT N application outweigh the costs of the investment in precision agriculture systems for cereal farms larger than 75 ha if basic systems costing £4 500 are purchased, and larger than 200 to 300 ha for more sophisticated systems costing between £11 500 and £16 000. Precise targeting of input applications within a field with VRT has the potential to improve input utilisation, increase input productivity and raise crop yields (Isik *et al.*, 1999).

In contrast, the study of Anselin, Bongiovanni and Lowenberg-DeBoer (2004), in which a comparison of the returns from different N application rates was made, modest results for VRT are shown. The study entails two uniform rates and a VRT. The first uniform N rate is lower, and recommended by agronomists. The second is higher, namely, the profit-maximising rate for the whole field, using the response functions. Returns above fertilizer costs vary with the N fertilizer rate recommended by agronomists, being the lowest at $415.35/ha, and the profit-maximising rate for the whole field is the highest at $419.56 /ha. VRT returns for recommendations by agronomists are higher than the uniform recommended rate, but lower than the whole-field profit-maximising rate, at $417.00 /ha.

Kahabka *et al.* (2004) also conclude that the results of their study indicate a low potential for increasing profits from site-specific N in that particular study field, which has high drainage variability.

In an experiment carried out by Welsh *et al.* (1999) in southern England to test variable-rate strategies, estimates of yield potential are produced from either historic yield data or tiller density maps derived from airborne digital photographs taken immediately prior to N application. Under a historic yield approach, three different strategies are investigated. It was found that, where 30% additional N is applied to a historically higher-yielding part of the experimental strip, yield is significantly increased by 0.54 ton/ha compared to the standard application rate. However, yield is significantly decreased by 0.48 ton/ha compared to the standard rate where N application is reduced in low-yielding areas. In cases where less N is applied to high-yielding areas, yield decreased by 0.9 ton/ha, whilst applying 30% more N to the low-yielding areas increased yield by 0.59 ton/ha. It is therefore apparent that applying more fertilizer to both the historically high- and low-yielding sections led to a significant increase in yield. Reduction in fertilizer rate, particularly in high-yielding areas results in a penalty of decreasing yield (Welsh *et al.*, 1999).

Under the tiller density approach, a small non-significant reduction in yield occurred where more fertilizer was applied to the high-density areas. The lack of a positive response is probably due to crop lodging. Reduction in fertilizer relative to the standard rate applied to a low-density area results in a larger yield reduction. Yield increase is also obtained where additional fertilizer is applied to the low tiller density area, as well as by reducing fertilizer application in high-density areas (Welsh *et al.*, 1999).

Malzer *et al.* (1999) observed that areas of the field that respond to higher N and P fertilizer rates to maximise returns in dollars, are not always associated with the highest marginal returns for that specific area of the field. This suggests a spatial difference in the efficiency with which the nutrients are used by the crop. Water/drainage/compaction differences within the study field, together with other non-detrimental field characteristics, have an influence on the yield, the yield response and the rate of nutrients required to obtain maximum economic returns (Malzer *et al.*, 1999).

Moss and Schmitz (1999) caution that, in view of the wide range of maize prices, N prices and agronomic differences in soil characteristics, investing in precision agriculture technology for the sole purpose of variable-rate fertilizer application in maize production is not profitable. Hurley *et al.* (2003) emphasise that, in determining the profitability of variable-rate fertilizer application, the cost of implementing a VRT strategy should be considered, as well as the potential for VRT to increase returns. Such costs may include the cost of collecting and processing the necessary information, as well as VRT equipment or custom application services (Hurley *et al.* 2003).

2.4.3.4 Variable-rate application of phosphorus

According to Clay *et al.* (2003), it is profitable to invest in VRT P management in some fields, but not in others. It is important to determine where VRT P management is profitable, and where it is not. Findings from the study of Clay *et al.* (2003) indicate that, if there are high to very high soil test P index ranges, precision P management may increase profitability. In this case, the crop must respond to added P fertilizer and the greater the P response, the greater the profit associated with precision P management. On the other hand, if low to medium soil test P index ranges are obtained, a relatively high P uniform application may be profitable (Clay *et al.*, 2003). This implies that the soil P index can influence the profitable application strategy, whether variable-rate or whole-field strategy. It is evident from the economic analysis that if the crop has a minimal fertilizer P response, then variable-rate fertilizer P management may not increase maize and soybean yields to an extent adequate to warrant grid sampling and the use of variable-rate equipment (Clay *et al.*, 2003).

In variable-rate application of input, interaction between different nutrients should be regarded in the same way as in the single-rate application of nutrients. Voortman and Brouwer (2001) contend that the response of yield to N depends on the available levels of other nutrients. They plotted the yield against N and P, and it was observed that the response to increasing N levels is almost positive, and virtually dependent on P levels. The return on N is highest at low N levels, and steadily decreases with increasing N. This means that, at low N, the additional units of N produce the highest yield increase, while the phenomenon of decreasing returns occurs as N levels increase. Yield increases due to N also strongly depend on the K level. A strong response is observed as N increases at low K levels, but no further response is detected at medium and higher K levels. The conclusion is that, at low levels, the response is strongest for each of the three macro-nutrients, but that the response decreases at higher levels, and may even be negative (Voortman & Brauwer, 2001).

2.4.3.5 Variable-rate herbicide application

Variable-rate herbicide application has the potential to reduce weed control costs and environmental loading of herbicides (Jensen & Hall, 2000). A small-scale research study conducted in Alberta, Canada, indicated that, in some cases, herbicide use can be reduced by 40% in certain field situations. However, the validity of these results at full-size field scale have not been established, as limited results that compare variable-rate and blanket spraying at field level, were obtained (Jensen & Hall, 2000).

2.4.3.6 Variable-rate seed application

Bullock *et al.* (1998) assert that variable-rate seeding of corn is of no economic value to farmers as a stand-alone. They conclude that this technology is only profitable to the farmer who is well-informed about the relationship between yield and seeding rate for each section of the field. Bullock *et al.* (1998) stress that, if this technology ever becomes profitable on a large scale and is adopted widely, it will become pertinent for farmers to know more about the characteristics of sections of their fields, how these characteristics vary, and how this variance causes yield response to seeding rates to differ across sections.

Lowenberg-DeBoer (1998), on the other hand, argues that variable-rate seeding is relatively inexpensive, especially for a producer who has already invested in GPS. Furthermore, larger gains can be expected in an integrated system that manages several inputs site-specifically. Economic advantages of variable-rate seeding have been observed on farms with some land that has a yield potential lower than 6 tons/ha. The proportion of low-potential soil determines the magnitude of the benefits, with benefits being highest when a small part of the farm has high-potential soil. When there is a small proportion of low-yield land, yield gains are the major source of the benefits, and benefits from seed saving are small. When there is a large proportion of low-yielding land, seed savings constitute the largest source of savings. Profitable variable-rate seeding is possible in mixes of high- and low-potential or medium- and low-potential soils. In cases of a mix of medium- and high-potential soils, single-rate seeding is better than variable-rate seeding (Lowenberg-DeBoer, 1998).

In conjunction with Lowenberg-DeBoer (1998), English, Roberts and Mahajanashetti. (2003) also conclude that the farmer's decision to adopt VRT depends on the proportion of soils with different potentials in the field. The spatial variability is very important in determining the extent of economic returns from VRT, and break-even budgets can be used for this purpose. The spatial break-even variability proportions are calculated as the proportions of different soils within a field with regard to which the use of VRT generates returns that are just adequate to recover the capital invested. To be able to make economic decisions, farmers must know the minimum spatial variability at which they can break even by using this technology. In their analysis, English *et al.* (2003) conclude that additional returns from the adoption of VRT more than cover the additional costs when the field consists of between 30% to 80% poor land.

2.4.3.7 Evaluation of other VRT applications

Jensen and Hall (2000) affirm that one of the real challenges in implementing variable-rate techniques stem from the fact that the development of GPS, GIS, computers, software, and field equipment has been faster than the agronomic understanding needed to bring together the enhanced information and capabilities provided by these technologies to the maximum benefit of the farm. It is possible for producers to apply herbicides and fertilizers at any point in the field, but the challenge lies in determining the specific rate for each location. The optimal amount of variable input depends on the yield price and variable input price, and – most importantly in the case of precision agriculture – on the level of fixed input, such as soil organic matter (Hurley *et al.*, 2003).

2.5 CONCLUSION

Precision agriculture, a crop production technique that leads to different management strategies in areas within a field that require different input levels, can be a key to success in farming. This technique is essential in sustainable farming in order to ensure sustainable use of resources, and to limit environmental pollution. Sustainability in agriculture can also be achieved by guaranteeing balances in the supply and removal of plant nutrients. Crop production functions assist in the economic analysis by specifying resource quantities and relationships between the inputs (supply) and outputs (removal). Precision agriculture encompasses a set of complex technologies that provide important information to assist in decision-making.

Profitability of precision agriculture, the single most important factor determining the adoption of precision agriculture technologies, has been evaluated by different researchers with regard to different inputs. The most commonly used strategies include the Ag, Ec and Inf strategies. These were evaluated relative to the whole-field management strategy. For lime, the net present value was higher and the crop response recovered precision agriculture costs, but this was not the case with potash, which resulted in negative Ec and Inf strategies. In the case of variable-rate seeding, profitability was high in soil mixes of high- and low-potential, or of medium- and low-potential soils. Some authors are in agreement on this point, while others contradict it. The value of switching over from whole-field management strategies to variable-rate strategies lies in correcting the misapplication of inputs, where some areas are under-fertilized while others are over-fertilized. The results are inconclusive, as some studies indicate profitability while others do not. Since the profitability of precision agriculture technologies and the determining factors may be site-specific, it is important to evaluate the profitability on a farm-by-farm basis.

Apart from the possibility of increased yields and reduced per-unit production costs that can be attained with precision agriculture, there are potential environmental benefits. Precision agriculture benefits are not without risk, and a wrong decision with regard to the technology can result in substantial risk. This risk arises mainly from yield variability and obsolescence. On the other hand, precision agriculture can be used as a risk management tool, since it provides more and better information.

The real challenge with precision agriculture is to identify factors that affect production in a given field, and to determine the appropriate strategy to maximise profitability for the farmer (Malzer *et al.*, 1999). Feasibility analysis determines whether the capital, labour, and other resources available are adequate to implement the proposed technology. The necessary management skills, adequate time, and sufficiently skilled labour are required to learn the technology and its associated problems. If any of these are lacking or inefficient, the adoption of precision agriculture technology is not feasible, even if it has been determined to be potentially profitable or otherwise beneficial (Lowenberg-DeBoer & Swinton, 1997).

As Anselin *et al.* (2004) state, the model specification used is crucial in the analysis of the profitability of precision agriculture, and it is pertinent to investigate which specifications produce reliable results. Models such as LRP, QPP and the QPR, as well as other non-parametric models, have been used in the past. In order to obtain more accurate estimates for the parameters in a yield response function, and ultimately for yield, return and profitability, proper data analysis techniques are required.

The experimental design, the methodology and data analysis technique used in this research will be discussed in Chapter 3.

Chapter 3

RESEARCH METHODOLOGY: EXPERIMENTAL RESEARCH

3.1 INTRODUCTION

Weiss (1996) defines agriculture as a phenomenon characterised by production that takes place over space. However, crop land is spatially heterogeneous in terms of yield, soil characteristics, landscape outline, pest populations and a host of other factors. The need for management practices based on, or suited to this spatial specificity, gives rise to precision agriculture. Precision agriculture makes it possible to collect, store, manipulate, analyse and ultimately act upon the huge amount of spatially localised information concerning the characteristics of a field.

Before turning to the methodological specifics of this study on precision agriculture, the main tenets of this research are conceptualised. As an on-farm comparison study, this research takes a practical approach by representing the actual farming operations. This was essential, as research is usually conducted by universities or commercial bodies in the form of "small plot" trials, but the adoption of recommendations based on such research is hampered by the fact that the conditions are considered too different from those of a real farm.

Since agriculture implies working a stretch of land over space, the orientation of this research is spatial. A spatial orientation acknowledges the fact that crop land is heterogeneous in terms of field characteristics. A model that provides guidance on the selection of characteristic variables in the analysis of precision agri,culture does not exist. The choice of variables is guided by factors of interest to the researcher or the problem at hand that needs to be solved. The most important variable in this study is yield, and its relationship with nitrogen (N) as an applied input. Lime is an important input in dry-land cash crop production in South Africa to correct the pH levels in the soil (Matela, 2001), which has an influence on the efficiency of other inputs. Besides lime, fertilizer is the most important input.

In order to systematise the concept of input variables, a distinction must be made between the independent and dependent variables. Variable-rate technology (VRT), which encompasses variable-rate application of N and the effective soil depth, which represents permanent soil characteristics, constituted the independent variables. Leedy (2001) explains an independent variable as a variable that the researcher manipulates. The independent variable in this study conformed to this definition since N was manipulated into different application rates as required by the soil potential, and treated as a variable that was considered the cause of variation in yield. Bailey (1994) indicates that, in an asymmetrical relationship, the variable capable of effecting change in the other variable is called the independent variable, while a symmetrical relationship implies that a change in one variable is accompanied by a change in another variable. The effects of variable-rate (VR) and single-rate (SR) treatments with regard to yield were the focus of measurement, and the effects of different strategies with regard to the managed independent variable, N, were also evaluated.

Maize yield was the dependent variable in this study, as it is potentially influenced by the independent variables and is, to some extent, dependent on them. The dependent variable can also be described as a variable of which the value is dependent upon the other variable, but which itself cannot affect the other variable. This is the case with fertilizer as, to a certain degree, yield level depends on the quantity of fertilizer (N) applied. All things being equal, a higher fertilizer rate increases yield to a certain extent. In a causal relationship, the cause is the independent variable and the effect is the dependent variable. The response of yield to effecting variables, VR applications of N and the effective soil depth were examined relative to SR application, which is essentially the control. The yield was measured as the effect or result.

A spatial orientation to the research methodology also has an impact on the choice of statistical methods. Weiss (1996) states that statistical analysis of subjects such as precision agriculture requires a spatially-oriented perspective. Many concepts traditionally described in economic theories and models in terms of a number, such as the total quantity of fertilizer applied to a farm field, now have to be described in terms of a surface, such as a VR fertilizer application surface of a farm field. This is the case when spatiality is taken into account. Numbers can be seen within a one-dimensional space, the number line. However, surfaces involve three-dimensional space, which increases complexity by involving two dimensions instead of only one. Statistical concepts such as the variability of yield over both space and time now need to be taken care of within this higher dimensional context. This kind of analysis demands the application of spatial statistics and spatial econometrics, rather than traditional time-series econometric methods. In the light of this information and the nature of the collected data, spatial econometrics was applied to examine relationships between yield and N treatments. When data has a spatial structure, more accurate estimates can be produced by spatial econometrics than by conventional econometrics (Bongiovanni & Lowenberg-DeBoer, 2001). Anselin (1999) explains spatial econometrics as a subfield of econometrics that attends to the treatment of spatial interaction (spatial auto-correlation) and spatial structure (spatial heterogeneity) in regression models for cross-sectional and panel data.

3.2 THE STUDY AREA

Before any management decisions can be applied to the farming situation, a detailed assessment of the land should be carried out. The natural abilities of the land should be fully assessed so that an overall picture can be formed. The farm studied is named Rietgat, and is located halfway between Wesselsbron and Bothaville districts in the Free State Province of South Africa. Rietgat is municipally demarcated under the Bothaville district. Bothaville is about 145 km southwest of Johannesburg. The farm is located approximately between Latitude -27E36N 5.9NN and -27E35N 30.9NN and Longitude 26° 32N 29.1NE and 26E32N 50.6N. On this farm, a 104 ha field was allocated to this study, with the following coordinates: Latitude -27E35N 57.6N and -27E36N 11.0N; and Longitude 26E33N 30.4N and 26E32N 38.0N.

The study period covered the years 2002/2003; 2003/2004; and 2004/2005. Initially, the study was to commence in 2001/2002, during which period phosphorus (P) was variably applied. However, P could not be variably applied in 2002/2003. In order to maintain consistency, the 2001/2002 data were therefore not used in the final analysis and writing up of the thesis. From the production year 2002/2003, only N was variably applied, and constituted the management variable.

Physical soil characteristics are important in determining the yield potential of a particular field. However, these depend on other important factors such as rainfall and other climatic factors.

3.2.1 Climatic conditions

All types of land use, especially crop production, are affected by climate. Climate determines the suitability of soils for crop production, and interacts very closely with the characteristics of the soil. In South Africa, successful cultivation of maize occurs in regions that receive an annual rainfall of 600 to 700 mm, of which at least 300 mm occurs through the months of December, January and February, the main growing period (AGR318, 1998). The average annual rainfall at Rietgat is 500 mm, which is adequate for the production of crops such as maize. The distribution of rainfall during the season is of greater importance than the total annual rainfall, which is not a very reliable measure. In years of good rainfall distribution, lower-potential soils can produce better yield improvement than higher-potential soils.

Maize germinates poorly at soil temperatures lower than 10 ^{o}C. Germination takes place between 10 ^{o}C and 13 ^{o}C; at this temperature level, fungal disease may affect the establishment of seedlings. A soil temperature above 13 ^{o}C is required for good germination and the development of strong seedlings. Temperatures during November and December, the period during which maize is planted at Rietgat, range from 15 ^{o}C to 30 ^{o}C, and are therefore suitable for germination. Day temperatures range from 22 ^{o}C to 32 ^{o}C, and are conducive to optimal growth of maize. As a hot-weather crop, maize requires day temperatures of at least 19 ^{o}C in summer for good production, while night temperatures should not fall below 13 ^{o}C. According to AGR318 (1998), the optimal growth temperature should be between 24 ^{o}C and 27 ^{o}C, depending on water provision.

3.2.2 Soil types

In order to ensure sustainability and long-term food security, it is imperative that soils are managed and utilized according to their potential to prevent degradation (Lal, 1991). South African soils are highly varied, not only between regions, but often also within the same field. Despite the fact that Bothaville soils are generally perceived as homogeneous soils, yields in this region show great variation. This superficial general opinion regarding the homogeneity of soils may therefore be a fallacy. The main spatial variant in this field is depth, as the soil survey revealed depth variation of a metre over a 200 m distance. A map of the various effective depths in the study field is shown in Figure 3.1 of Annexure 3, while Figure 3.2 of the same annexure depicts a topography map indicating different drainage capabilities in various parts of the field.

The dominant soil form in this field is Avalon (70.80 ha), followed by Pinedene (25.75 ha), with tiny spots of Westleigh, Clovelly, Molopo and Sepane. Under the world reference base, the dominant soil form is classified as Eutric Arenosol. Soil forms are the result of the unique soil-forming factors and processes that took place in a specific soil profile. Soil forms are defined by grouping together diagnostic horizons (Soil Classification Working Group, 1991). A soil map that depicts different soil forms within the field is attached as Figure 3.3 in Annexure 3.

Soil texture and depth are the basic properties of the soil, as they do not readily change. The effect of texture on the soil water properties and chemical properties of the soil brings its importance to the fore. The clay percentage of soils on the study farm ranges from 6 to 15%. The soils on the ridge are sandy Avalons with an estimated 5% clay in the topsoil, 10% clay in the subsoil and 25 % clay in the soft plinthic B-horizon. The soil type changes to Pinedenes lying in the lower part. The clay content is also very low in the topsoil (5 %), but about 25 % in the yellow-brown apedal B-horizon. The transition is an Avalon with slightly higher clay content in the yellow-brown apedal B-horizon. The Pinedene is very dark in colour.

Le Roux *et al.* (1999) refer to the effective soil depth as the total thickness of the A- and B - horizons, or solum. The effective soil depth determines the volume of the soil, as well as the amount of moisture that can be stored. The depth limit of the South African Soil Classification is 1.5 m on diagnostic horizons. Soil depth is important in the semi-arid regions as it determines the water retention and thus the danger of water-logging or drought, and therefore the production potential of the soil. The soil depth changes from >1 200 mm in sandy Avalons on the ridge to 800 mm in Pinedenes lying in the lower part. The limiting layer in a soil profile, such as an unspecified layer with signs of wetness in a Pinedene soil, has an advantage for the precarious rainfall in the semi-arid regions in the sense that a greater percentage of the profile water is consequently available to plants (Le Roux *et al*., 1999).

The average depth of 1 500 mm in the study field is better than the average optimum depth of 800 mm in the Western Free State (Le Roux *et al.*, 1999). Effective depth is important in evaluating the suitability of the soil, but not in the description or interpretation of the soil. Maize requires an effective depth of at least1 200 mm, and the effective depth of 1 500 mm in the study field meets this requirement.

The sandy Avalon soil has a very high potential for maize production. It has a high water infiltration rate and water storage capacity, and stores its water deep. The transitional area of Avalon is restricted by a shallower depth, which may inhibit maize production when poorly-distributed rainfall occurs. However, over a period of several years with well-distributed rain, it may produce the same as the high-potential soils. The Pinedene soils are very wet soil types, and production is restricted by excess water in the rainy season. These soils are very wet, and the morphology indicates wetness rather similar to much wetter soils. In very dry years, these soils may outperform the soil types that are generally considered “good” maize soils.

Attention will now be given to the methodological specifics of this chapter, which discusses the methods used in gathering and analysing data. The focus is also placed on the experimental design used and the general philosophy of on-farm research design. The reasons for selecting the design used are also explained. This is followed by the experimental procedures and techniques used to gather data in the field. Data analysis methods, model specification and techniques used in estimating the model are also discussed. The main method of data analysis, spatial econometrics, is explored together with the estimation thereof and diagnostic tests for the spatial dependence that may be present in the data.

3.3 EXPERIMENTAL DESIGN

The strip-plot design (Brouder & Nielsen, 2000) was used for this on-farm research experiment. In on-farm experiments where small plots are impractical, strip plots are recommended in order to adjust to the classical agronomic statistics. The design used involved treatments that run to and fro across the field, corresponding to planting and harvesting. Two treatments were investigated – the VR application of N and SR application strategy of the same input. According to Safo-Kantanka (1994), this design is used when known differences between the experimental units exist, and may contribute to the variation between the tested factors. There are spatial gradients in the effective depth levels in the experimental field, as revealed by an intensive soil survey, and this variable could affect yield response to different treatments. If this factor was not taken into consideration, it could have bloated the experimental error.

3.3.1 On-farm comparisons

Before venturing into the specifics of the experimental design of this study, it is imperative to highlight some important issues relating to on-farm research. On-farm comparisons of new technologies are often conducted by farmers in large non-replicated blocks, or split-field comparisons. Although the validity of these methods is questionable, especially with regard to the traditional agronomic research methods, the information generated is useful for management decision-making.

Major changes in weather patterns and technology occur on a yearly basis, and differences in soils within and among the fields are also apparent. In practice, these factors may compel a farmer to make adjustments along the way, and even to modify the on-farm research goal. If modifications are made, it is essential that the trials are repeated for a number of years, or even on a number of fields (Nafziger, 2002).

Nevertheless, these on-farm comparisons with inadequate replication data can be useful in the sense that they seize opportunities created by precision agriculture technologies and spatial regression methods. With the aid of global positioning systems, VRT equipment, scouting tools and the related software, important information from field-scale plots can be captured (Boyer, Erickson & Hawkins, 2005). A yield-mapped field can generate large amounts of site-specific response data in on-farm experiments (Pringle, Cook & McBratney, 2004), and spatial analysis can help explain spatial patterns in soil attributes and yields (Griffin, Lambert & Lowenberg-DeBoer, 2004a; Griffin *et al.*, 2005b). Griffin *et al*. (2005b) assert that appropriate analysis of the systematic aspect of the variability can increase confidence in farmers' experimental results. Furthermore, the innate spatially-correlated yield monitor data can be modelled with spatial regression methods.

In recent years, site-specific crop management (SSCM) on-farm research has also gained momentum. SSCM on-farm research differs from traditional on-farm research. Firstly, traditional on-farm research does not take spatial and temporal variation in agronomic data into account. Secondly, SSCM on-farm research implies that large quantities of data that were not available before are now accessible, but with analysis implications. Lastly, the fundamental nature of SSCM on-farm research is such that it causes minimal interference with the farmer's normal cropping operations, making simplicity, flexibility and risk management of paramount importance (Pringle *et al*., 2004).

Mallarino (1998) alleges that, by conducting on-farm strip trials, precision agriculture technologies can contribute to evaluating and demonstrating alternative fertilization or other management practices, and that this methodology is useful for new research, adjusting recommendations to local conditions, and understanding the causes of yield variability better. Griffin *et al.* (2005b) found that local variation over the production surface can be reflected more accurately when spatial auto-correlation is taken into account in the analysis of yield monitor data. In the case of Monte Carlo simulations, Griffin *et al.* (2004b) discovered that the accuracy of analysis of variance (ANOVA) decreases when classical assumptions are made, while spatial ANOVA makes it possible to correctly estimate the true model parameters at 5%.

Griffin *et al*. (2004b) point out an important issue related to on-farm comparison. The objectives of farmers and researchers are often different. Researchers are mostly interested in producing generalisable results, while farmers' interests lie in the improvement of management decisions, and ultimately profit. The time and energy invested by farmers in on-farm research are only justified if the benefits arising from the new knowledge are more valuable than the incurred costs (Lowenberg-DeBoer, 2002). It is also important to remember that no general recommendation made on the basis of on-farm research is valid for all fields on a specific farm (Mallarino, 1998).

3.3.2 Plot layout

It was deemed important that the plot layout used can be employed by farmers to determine an optimal N application strategy in any particular field. To achieve this, it was necessary to use standard farm machinery for the experiment and move away from the traditional small-plot randomised block experimental design. According to Griffin *et al.* (2004b), the eventual test of an experimental design in on-farm comparison is whether it facilitates better decision-making. Small-plot designs in traditional agronomic trials reduce heterogeneity, but require intensive planning, management and labour input during planting and harvesting. They are also cost-intensive, and interfere with regular farming activities. Treatment interactions due to proximity of treatments constitute another disadvantage of small-plot designs.

Large experimental units (>5 ha) can solve some of the problems associated with small-plot designs for on-farm comparisons. Interference with farming activities is limited, and the designs can be easily implemented by farmers. Large experimental blocks have other benefits at farm level. Geo-referenced information per hectare is provided by precision agriculture technologies such as Global Positioning System (GPS), yield monitor and proximate sensors, at a fairly low cost. The farmer's input at harvesting time does not increase substantially, except with regard to calibrating the implements to be used. The ease of data gathering with the available technology motivates farmers to conduct on-farm comparisons. Lastly, treatment edge effects can be managed effectively in a large experimental unit, since dubious points can be deleted without compromising data availability (Griffin *et al.*, 2004b).

The plot layout for this field trial is depicted in Figure 3.1.

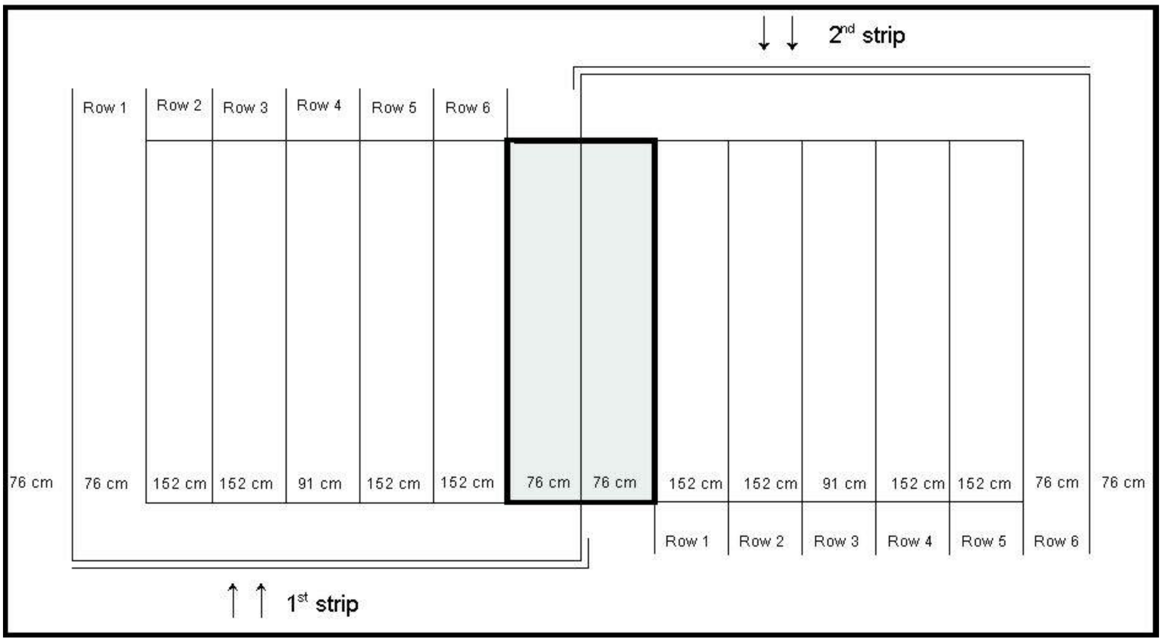

Figure 3.1: Experimental design: Strip-plot layout

A 104 ha field on the farm Rietgat was divided equally into strips, with six rows for VR application, alternating with six rows for SR application of N (Figure 3.1). Each set of six rows (strip) constituted a block, or plot. Each treatment strip in the field ran across good, average and poor zones, i.e. crossed over different potential zones (full details provided under the experimental procedure and techniques section). This design was used to compare fertilizer application programmes under VR and SR applications. As the planting was done back and forth across the field, multiple random side-by-side replicates were obtained. Figure 3.2 depicts a map of management zones identified on this field.

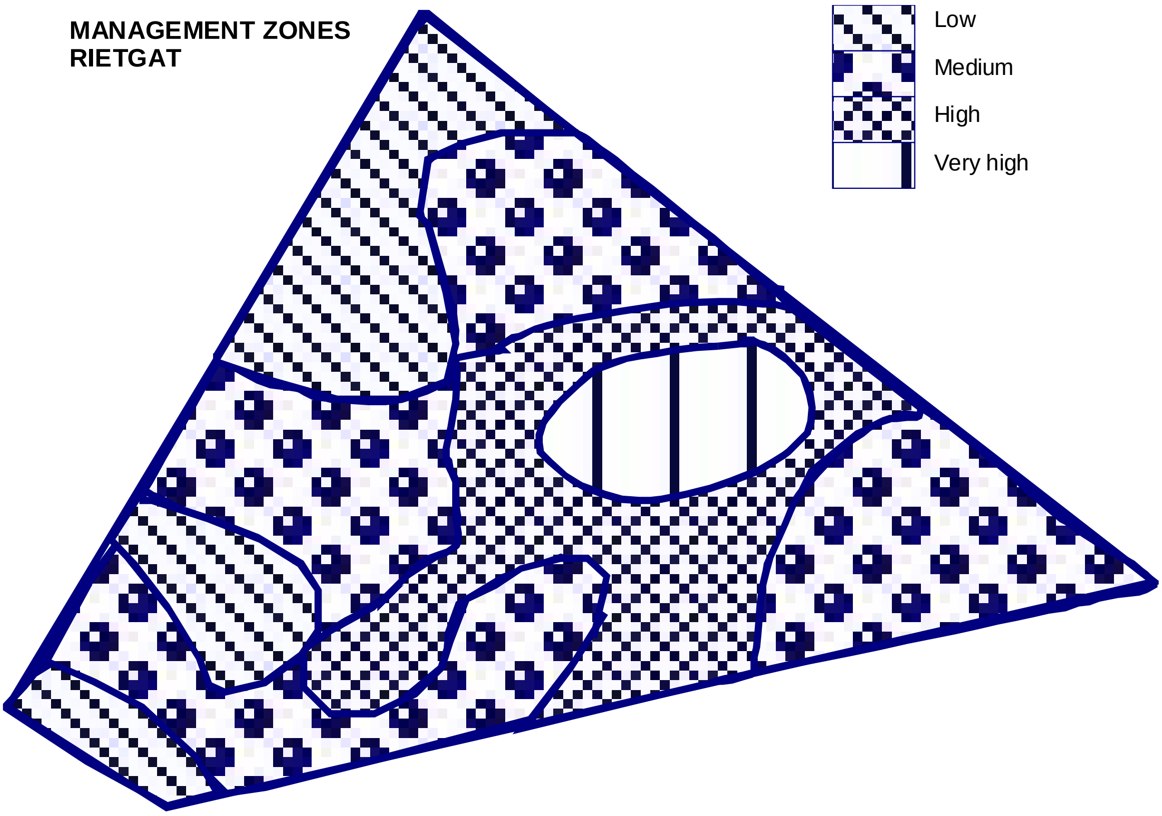

Figure 3.2: A map of management zones

Brouder and Nielsen (2000) suggest that the width of the strip should be a multiple of the width of the combine head. The widths of the strips (8.5 m) in this study were equal to the width of the fertilizer applicator, planter and combine harvester, and maintaining consistency in strip widths facilitated field operations. According to Brouder and Nielsen (2000), it is best to have the treatment run the full length of the field to avoid the need for adjustments part of the way through the trip across the field. It is recommended that the length of the plot (strip) be a minimum of 350 feet (107 m) to allow removal of end rows, and to ensure that the combine can maintain a constant speed, as required by the yield monitor, for the greater part of the trip or pass (Brouder & Nielsen, 2000). The treatments in this study satisfied these conditions, as the minimum length was the recommended 107 m.

It is appropriate to acknowledge the limitations of strip-plot designs. Apart from the problem of spatial effects, strip-trial designs allow spatial edge effects at treatment borders. Such spill-over effects influence the measurement of spatial variability, thereby affecting the power of estimating treatment differences. This is what Griffin, Lowenberg-DeBoer and Florax (2006) call the "neighbouring observation problem".

3.3.3 Replication

Wittig and Wicks (1999) explain replication as an important component of on-farm comparison. When two treatments are compared, it must be ensured that each is independently represented a number of times, and this can best be accomplished by the use of alternating strips based on equipment size. On this trial, 160 replications were done. Replication was important to ensure that the results represent a response to a VR versus an SR treatment, and not just a random variation that occurs naturally across the field. Bailey (1994) asserts that, on average, three replicates are sufficient for a meaningful analysis, but having four or more is optimal. In this study, where expected differences between soil types are small, more replicates were needed. Having more comparisons provided a better idea of the size/magnitude of the effect of treatments (VR and SR). Mallarino (1998) emphasizes that a valid conclusion concerning differences between treatments applied to strips can only be reached if treatments are replicated across the field, independent of the length or the width of the strip used. Comparison between the two treatments made it possible to ascertain the resulting change in costs, yield and consequently profit, emanating from each treatment. This information on the size of the response to the treatment allowed an evaluation of the profitability, which could otherwise not have been made validly. The replication treatments were identical to the initial treatments, and were carried out on the same location in each season of the three-year maize production period. This means that strips of VR and SR applications were identical each year. However, caution should be exercised with narrow strips of one to three combine passes. Narrow strips can allow treatments such as herbicides and fertilizer to drift to adjacent strips or move out of the experimental plot, interfering with other treatments. The fact that this can sometimes result in less reliable analysis (Griffin *et al*., 2004a), should be taken into account.

Although it is essential, replication alone – without cognisance of spatial auto-correlation – cannot improve the reliability of standard designs. Griffin *et al.* (2004a) found that OLS and spatially-generalized least squares (SGLS) resulted in similar estimates in the absence of spatial auto-correlation, using the Means Square Errors (MSE) as a measure of good estimates. However, as spatial auto-correlation increases, MSE increases in both models, but to a greater extent in the OLS model. On the other hand, MSE decreases as the number of replications increase, indicating the importance of both replication and accounting for spatial auto-correlation in analysing spatially auto-correlated yield monitor data.

3.3.4 Randomisation

The strip-plot design that was used, allowed randomisation of treatments to ensure an equal probability for each treatment to be affected by both known and unknown spatial differences in the experimental site. The back-and-forth movements of the equipment creating treatment strips ensured randomisation, and no prior deliberations were made regarding the treatment allocation in terms of particular soil characteristics. Safo-Kantanka (1994) suggests that, instead of just randomly distributing the treatments to the experimental units or plots, the experimental units can be grouped into blocks and the treatments be assigned to the experimental units or plots within a block. This was accomplished by grouping six rows together to form a block (strip), which can also be referred to as a replicate. Each strip was therefore considered a complete experiment to which a certain fertilizer application strategy (VR or SR) was assigned.

Brouder and Nielsen (2000) warn that a completely random assignment of treatments to plots or strips can result in some bias arising from the side-by-side comparison. For instance, it could happen purely by chance that the same treatment (VR) is assigned to plots on the upper part of the slope more than once. This presents a risk that treatment effects may not be easily distinguishable from topographic effects. As was the case in this research study, it is best to use a design that randomises treatments between replicates. Since the experimental field is relatively flat, topographic effects were minimal, but permanent soil properties (the effective soil depth and the clay percentage), as well as applied inputs, had an effect on the variation in yield. Given that treatments were allocated randomly between and not within strips, extra error correlation was anticipated, which could take the form of treatment strip correlation in the same way as in the study by Hurley, Malzer and Kilian (2003). As a result, treatment and treatment strip heteroscedasticity and spatial auto-correlation had to be accounted for in order to obtain a precise estimation of yield changes and resulting profit.

3.4 THE EXPERIMENTAL PROCEDURES AND TECHNIQUES

The John Deere Office Software System enabled identification of soil sampling points within management zones as identified by a coordinates system, and the amount of each nutrient that had to be applied was determined on the basis of the fertilizer guidelines provided by the National Fertilizer Society of South Africa (NFSA), included as Table 3.1 in Annexure 3. With the aid of the John Deere Office Software, the input applications were continuously varied throughout the field, as determined by the prescription card set up beforehand by the farmer.

3.4.1 Field Management

Field management encompasses all field activities, beginning with soil preparation and ending with harvesting. Of importance in this study are the cropping systems and the procedures followed in fertilizer application. It should be pointed out in passing that low pH problems were identified on about 40 ha of the field, and lime was applied in the first year of study (2002/2003) to rectify the pH in this area and bring it to the acceptable crop production level. This is the essence of on-farm comparisons, as problems experienced are corrected during the season. During the data analysis, it became evident that the lime application had a tremendous effect on the management zones. Zones that were regarded as low-potential did not produce low yields, as had been expected. A zone that was initially classified as low-potential, produced yields higher than those obtained in the medium-potential zone, and performed just as well as the very high-potential zone. It emerges that the acidity status of the soil contributed to the low productivity of the soil in certain parts of the field, rendering them low-potential. The application of lime changed the productivity of such areas.

3.4.1.1 Soil preparation

The cropping system involved the production of conventional white maize, which has been produced in this field each year for the past 10 years, coupled with a winter fallow period. Maize is planted in November / December, and harvested in June / July. Different management activities are carried out in the field throughout the season. Shortly after harvesting the proceeding crop, a V-blade implement that is about 15 cm deep, is used to cultivate the field in an endeavour to kill winter weeds. This is coupled with P application in the form of super-phosphate and row marking, since the same tracks are used year after year for the movement of implements.

In August/September, deep ripping to a depth of 700 to 800 cm is performed on rows to allow deep root and moisture penetration. The depth depends on the soil texture, as it is difficult to rip down to 800 cm on some soils. This deep cultivation is essential, since shallow cultivation can limit the realisation of the soil's yield potential. Depending on the levels of weed infestation in the field, herbicide (Roundup) is sometimes applied at the same time. About one month before planting, the final tillage operation is carried out and N fertiliser in the form of urea is also band-placed on rows, about 100 mm deep and 100 mm away from where the seed will be planted. Cleaning cultivation is done during planting with a rolling rod implement that turns the soil to a depth of about 10 cm. This cultivation activity also depends on weed infestation, and is not implemented if there is no weed problem. A nitrogen fertiliser mixture is also applied during planting. Four to five weeks after planting, N top-dressing is banded as liquid Urea Ammonium Nitrate [UAN (32%)], deposited on top of the soil and worked in slightly. Pesticide and herbicide applications are done during the season as the need arises.

3.4.1.2 Nitrogen application

The aim of the fertilizer application technique chosen was to allow investigation into the effect of VR N application on yield, in comparison with the SR application method. In the case of VR, N application was based on the yield potential of the soil, as classified into different management zones. Yield potential, deducted from historical yields, served as an attribute that identified management zones and varied N rates for this study. Average yield was used as a determinant of yield potential in O'Neal *et al.* (2004). While O'Neal *et al*. (2004) estimated yield potential as an average of the yield monitor data plus 5% for the whole field, as well as for each grid cell, a simple average yield monitor data for each management zone was used in determining the yield potential per zone for this study.

The process of identifying management zones for N application involved layering yield maps of the past three years. Average yields were established by means of this layering, and an average yield map was hand-drawn by the farmer on the basis of the layered maps. A polygon was created around the soil survey data, and the yield data was appended to the soil survey data. Data from the other two years were appended to the same soil data using ArcView GIS 3.2, in order to ensure that the yield maps for three years had the same resolution. This yield map was to form the basis for digitising the hand-drawn yield map. The idea was that the best and worst years could be discarded over a 10-year period, but with only three years of spatial data available, it was more practical to use averages.

Four management zones were identified. First was the low-potential zone, with a target yield of less than 3 tons/ha. The medium-potential zone, with a potential yield of 3 to 4 tons/ha, constituted the second management zone. Following this was the high-potential zone, established with a target yield of between 4 and 5 tons/ha, and finally the very high-potential zone, which had a potential yield of more than 5 tons/ha. The yield map was compared with the soil type map to identify any correlation. There was a high correlation between the two maps regarding potential levels, and the yield map was subsequently used to create a fertilizer application map. The demarcation of management zones primarily used historical yield data (a rough guide at best), but it later emerged that the acidity problem in the soil confounded the allocation of management zones.

The yield potential of different zones was taken into account in determining fertilizer rates. N tests were not performed as the soil is sandy, and it was assumed that all applied N would be used up by plants and the rest would be leached and/or volatilised. As a result, the N status in the soil was not taken into account, but the target yield as indicated in the fertilizer guidelines handbook provided by NFSA (Table 3.1, Annexure 3) was used as a guideline. A six-row John Deere fertilizer applicator was used, with 152 cm width between the two extreme rows on both sides and 91 cm between the two middle rows, through which the tractor passes. This is illustrated in Figure 3.1. The row widths were five feet (1.5 m) and three feet (0.9 m), with an average of 1.41 m. The longest row measured 1 502 m.

Depending on soil moisture conditions, the effect of N on yield varies from season to season. Based on the soil potential, yield increases if enough N is available in the soil during rainy seasons; in low rainfall seasons, however, high N leads to a decline in yield.

Because of this relationship between N, yield and soil moisture, the expected rainfall influenced the amount of N applied. As the total amount of N applied depended on the amount of rainfall (moisture) available during the growth season, it was strategically important that not all the N needed for the whole season was applied during planting. N top-dressing in the form of liquid N was implemented as a management strategy, in which potential yield was controlled on the basis of the available water during the growing season.

In the first year (2002/2003), a 5:2:1 (32) mixture (20 kg N, 8 kg P and 4 kg K per 100 kg fertilizer mixture) was band-placed before planting, while variable rates of N as top-dressing in the form of UAN were band-placed. Four weeks before planting, a constant rate of 30 kg N per hectare was applied throughout the whole field in the form of urea (46%). The second application of N, still constant for the whole field, was done during planting, with a flat-rate mixture of 5:2:1 (32) at 30 kg N per hectare. The after-planting total amount of N applied amounted to 60 kg N per hectare [30 kg from urea and 30 kg from 5:2:1 (32)].

The totals for each nutrient were as follows after planting:

N: 30 + 30 = 60 kg N per hectare

P: Various rates 18 + 12 = 30 kg P per hectare

K: 6 kg K per hectare

This represents the base application for the whole field, without differentiating between VR and SR treatments. Application strategies for a top-dressing differed between the two treatments (VR and SR). Band-placed top-dressing with UAN (32%) was done five weeks after planting according to the yield potential and was map-based, with variations of 24, 32, 40 and 50 kg/ha. The total N applied for the whole season ranged from 84 to 110 kg/ha. The top-dressing rate for the constant application was 45 kg N per hectare.

A 5:2:1 (32) mixture proved too expensive in the second year (2003/2004), and the farmer switched to a 4:2:1 (28) mixture (16 kg N, 8 kg P and 4 kg K per 100 kg fertilizer mixture).

It was established that continuing to use urea for top-dressing and substituting 5:2:1 (32) for 4:2:1 (28) would reduce costs, but still meet the nutrient requirements of the soil. A constant rate of 36 kg/ha of N in the form of urea was applied on the entire field three days before planting. This urea should have been applied four weeks before planting, but was postponed until three days before planting due to drought conditions, in the hope that it would rain in time before planting. Another 24 kg/ha of N was applied during planting.

For VR application during top-dressing, N in the form of UAN was varied in four application rates of 9 kg/ha for the low-potential zone, 25 kg/ha for the medium-potential zone, 50 kg/ha in the case of the high-potential zone, and 63 kg/ha for the very high-potential zone. Approximately 42 kg/ha UAN (32%) was used for top-dressing in the SR application treatment, all band-placed as in the first year. The total amount of N applied in the second year ranged from 69 to 123 kg/ha in the VR treatment.

In the third year (2004/2005), a 4:2:1 (28) mixture was still applied on the entire field before planting at a constant rate of 60 kg N per hectare. UAN top-dressing was band-placed five weeks after planting in a variable manner. Four rates of 9 kg/ha, 25 kg/ha, 54 kg/ha and 67 kg/ha were applied in the VR treatments, while a constant rate of 42 kg/ha was applied in the SR treatment. Seasonal N application in the third year varied from 69 to 127 kg/ha.

Generally, K deficiencies seldom occur in the maize-producing areas of South Africa, and K requirements therefore took a back seat. A summary of N application rates for all the three years is provided in Table 3.1.

TABLE 3.1: NITROGEN APPLICATIONS

Nitrogen UAN 32% (kg/ha)	2002/2003		2003/2004		2004/2005	
	VR	SR	VR	SR	VR	SR
Rate 1	84	105	69	102	69	102
Rate 2	92		85		85	
Rate 3	100		110		114	
Rate 4	110		123		127	

The rates encompass total seasonal N application.

3.5 DATA

3.5.1 Data collection

Yield data was collected with a combine harvester fitted with a yield monitor that has reference to location, with yield measured as points in space. Yield was recorded on a yield monitor card by a yield monitor equipped with a GPS, which indicated yield at a specific place (height, latitude and longitude), as well as moisture content of the grain. Data regarding the speed of the combine harvester, the harvesting rate and the time in seconds were also recorded during harvesting.

Before the yield data was analysed, it was edited, filtered and cleaned using Yield Editor software (Drummond, 2005). Erroneous observations were identified and removed by looking at harvester velocity (minimum and maximum); velocity change, start- and end-pass delay, maximum and minimum yield, flow delay and the standard deviation. The measures used for these variables are presented in Table 3.2.

TABLE 3.2: DATA-FILTERING FACTORS AND CRITERIA

Factor	Standard
Maximum velocity (km/hour)	5
Minimum velocity (km/hour)	1
Velocity change (km/hour)	1
Start-pass delay (seconds)	5
End-pass delay (seconds)	5
Maximum yields (tons/ha)	10
Minimum yields (tons/ha)	1
Flow delay (seconds)	5
Standard deviation	3

These parameters were specified, and the data was edited, filtered and cleaned in a matter of seconds.

The collected spatial yield data was processed into maps to reflect variations that exist in the harvested yield. When mapping a set of spatial data observations, two sources of information are used for quantifying locations. The first source of information is derived from the Cartesian space that is represented by latitude and longitude. On the basis of this information, distances from any point in space, or the distances from observations located at distinct points in space to observations at other locations, can be calculated. To fulfil this condition, yield data was collected with the yield monitor, which recorded yield and the associated latitude and longitude, thus making it possible to quantify the location using Cartesian space. According to LeSage (1998), spatial auto-correlation should conform to the fundamental theorem of regional science, which states that distance matters. A greater degree of spatial auto-correlation should be reflected in observations that are located in close proximity (within the same management zone) than those separated by greater distances (between management zones). Therefore, the strength of spatial dependence between observations declines as the distance between observations increases.

Contiguity (neighbouring), which reflects the relative position in space of one regional unit of observation in respect of other such units, is the second source of locational information.

Lambert, Lowenberg-DeBoer and Bongiovanni (2004) define contiguity as a function of the distance that separates one grid or polygon from another. Blocks or zones in the same locality share a similar weight, and the spatial weight matrix is described by the combination of neighbourhoods covering the entire grid. Knowledge of the size and shape of the observational units depicted on a map determines the measures of contiguity. It can be ascertained from this which units are neighbours and have borders that touch, or represent observational units located in reasonable proximity to one another. There should be a higher degree of spatial dependence in neighbouring units than in units situated far apart (LeSage, 1998). However, this method of quantification was not used in this study.

LeSage (1998) asserts that quantification of the locational aspects of the sample data is much more important than inquisitions into spatial auto-correlation and spatial heterogeneity. This quantification is essential in the computation of spatial weight matrices to describe spatial relationships between observations when data are analysed.

3.5.2 Data analysis

As the data collected were not limited to yield monitor data only, all data (soil and application data) had to be assembled together. ArcView GIS 3.2 software was used for this purpose. Dummy variables representing the two treatments were allocated to the yield and N application data in Excel before assimilating the data in ArcView GIS 3.2. As the soil data represented the least dense data, a seven-metre buffer was created around the soil point, and the yield points within this buffer were averaged and assigned to the soil point. The N application data was also appended to the buffer. The size of the buffer was selected based on the size of the treatment strip, which was 8.0 m. The fact that the buffer distance of seven metres was slightly less than the strips of different treatments ensured that yield data from adjacent treatments were not averaged together.

The collected data were analysed in three phases by means of exploratory, statistical and profitability analysis methods. Exploratory data analysis is an approach that draws on graphical techniques to maximize insight into a data set by uncovering underlying structure, extracting important variables, detecting outliers and testing underlying assumptions regarding anomalies. This graphical visualisation makes it possible to gain insight into the characteristics of the data before conducting any complex and detailed analysis. A detailed discussion of this method and the results obtained is presented in Chapter 4.

As in the case of the study by Bongiovanni and Lowenberg-DeBoer (2001), spatial econometrics in this study was concerned with estimating the relationships between variables that have spatial structure, such as soil nutrient N and soil attribute (effective depth), which are then used to calculate outcomes of economic interest, such as yields, profits and costs. These outcomes are an important tool in management decision-making, particularly with regard to the implementation of VRT to manage variability (Bongiovanni & Lowenberg-DeBoer, 2001). The statistical and profitability analyses are presented in Chapter 5.

Correlation among neighbouring observations, as often found in yield monitor data and other spatial data, violates the assumptions of classical statistical analysis, particularly with regard to homoscedasticity and the absence of auto-correlation (Lambert *et al*., 2004). This makes analysis of this type of data rather difficult and invalid from the viewpoint of classical agronomic research, as the ignorance of spatial structure results in variance estimates that tend to be inflated and the significant levels of test statistics therefore tend to decrease, resulting in unreliable statistical inference. Due to spatially auto-correlated data, the use of statistical tools such as Ordinary Least Squares (OLS) to analyse crop response functions can lead to inaccurate OLS estimates, which result in an under-estimation of heterogeneity and inefficient or biased inferences. Imprecise inferences can thus be made about the profitability analysis of trials comparing VR to SR application rates of N (Lambert *et al.*, 2003).

In the following subsections, a short description is provided regarding the two spatial effects, spatial auto-correlation and spatial heterogeneity, which have to be taken into consideration in analysing precision agriculture data. A discussion of diagnostic tests for these spatial effects follows this description.

3.5.2.1 Spatial auto-correlation

Bongiovanni and Lowenberg-DeBoer (2001) describe spatial auto-correlation as the situation where the dependent variable or error term at each location is correlated with observation on the dependent variable or values for the error term at other locations. Spatial auto-correlation can be formally expressed by the moment condition:

$$Cov\ [y_i, y_j] = \varepsilon\ [y_i\ y_j] = \varepsilon\ [Y_i].\ \varepsilon\ [Y_j] = 0$$

$$for\ i \neq j$$

where $\boldsymbol{i}$, $\boldsymbol{j}$, refer to individual observations (locations) and y_i (y_j) is the value of a random variable of interest at that location.

LeSage (1998) elucidates further that, in a collection of sample data observations, spatial auto-correlation refers to the fact that one observation associated with a location, which may be labelled i, depends on the other observation at location $j \neq i$. This is formally stated as:

$$y_i = f(y_j),\ i=1,......., n \qquad j \neq i$$

In yield data, spatial auto-correlation is caused by coincidence of similarities between location and values obtained, such as yield. This can be ascribed to the fact that there is always a high chance that high or low values for a random variable will be surrounded by neighbour observations with similar values. Since the field was divided into different management zones where areas of equal yield potential were grouped together, auto-correlation was anticipated. As a result, a yield point at a specific point was expected to correlate with a yield point at another point within a management zone – thus the inevitable presence of auto-correlation. The presence of auto-correlation was determined in the GeoDa™ statistical package for spatial data analysis (Spatial Analysis Laboratory, 2004) using the Moran's I statistic, the Likelihood Ratio and the Lambda Correlation Coefficient.

3.5.2.2 Spatial heterogeneity (Heteroscedasticity)

LeSage (1998) explains the term spatial heterogeneity as variation in average relationships between X and Y over space. This implies that a different relationship may be expected for every point in space. The relationship between *N* as an *X* variable and yield (*Y*) varied from one point to the next, or from one management zone to the next. LeSage (1998) states that, when sample data is associated with a location, spatial dependence exists between the observations, and spatial heterogeneity occurs in the relationships being modelled. Bongiovanni and Lowenberg-DeBoer (2001) refer to heteroscedasticity as a case in which the variance of the error term is not constant for all values in the independent variable, with a resulting unequal reliability between Y values. In a linear relationship, spatial heterogeneity can be expressed as follows:

$$y_i = X_i \beta_i + \varepsilon_i$$

where *i* indicates observations collected at spatially-referenced management zones (i=1,...,n). X_i represents a (1 x k) vector of explanatory variables (N) with an associated set of parameters (β_i), y_i is the dependent variable, yield at location *i*, and ε_i denotes a stochastic disturbance in a linear relationship and captures heteroscedasticity. The presence of spatial heterogeneity was tested by the Breusch-Pegan (BP) and Koenker-Bassett (KB) tests in the GeoDa™ statistical package. It is important that thorough diagnostic analysis of the two spatial effects is conducted, so that the appropriate model that accounts for these effects is selected.

3.5.3 Diagnostic tests for spatial dependence effects

Diagnostic tests for spatial dependence in the OLS model helped confirm the presence of spatial auto-correlation. The Durbin-Watson *d*-statistic, the most celebrated and most commonly used test to detect auto-correlation, was used as a starting point to detect the presence of any auto-correlation. However, this method is mainly used for first-order auto-correlation to compare the residual for the time period *t* with the residual from time period *t*-1 (Gujarati, 1999). This method is not appropriate for testing higher-order correlation such as correlation in space.

According to Bartlein (2003), the first step in conducting a spatial regression analysis is defining the elements of the spatial weight matrix, *W*, that describe the spatial relationships (distance, contiguity, etc.) between the observation points. This is because the presence of spatial effects can be easily determined when a regression model is estimated together with the spatial matrix linked to it (Lambert *et al.*, 2004). The potential spatial interaction between two observations is indicated by non-zero elements.

In testing for spatial effects (spatial auto-correlation and heteroscedasticity) in econometric models, spatial weight matrices had to be constructed and then included in a specified regression model in order to control for some of the spatial effects. The incorporation of the spatial weight matrix into the regression determines the relations between the dependent variable y_i and neighbouring y_j's. Spatial weight matrices were constructed in GeoDa™ using the minimum Cartesian distance, which ensures that each observation has *at least* one neighbour, and included in a Baseline Model and numerous Sensitivity Test Models to test for spatial effects. Moran’s I and Lagrange Multiplier tests were used mainly to detect spatial auto-correlation, while BP and KB tests were used largely to detect heteroscedasticity.

3.5.3.1 Moran’s I test

Moran’s I, the most commonly used specification test for spatial auto-correlation (Anselin & Florax, 1995), was used for detecting spatial auto-correlation in the data collected. Moran's I measures spatial auto-correlation in regression residuals and indicates the degree of spatial association between the yield points for each of the weight matrices (Bongiovanni & Lowenberg-DeBoer, 2001). The Moran’s I statistic is quite similar to the popular Durbin-Watson test, and is formally equivalent to the slope coefficient in a linear regression of *Wx* on *x*, where *x* is the standardised value of yield (Bongiovanni & Lowenberg-DeBoer, 2001). In addition to the computed Moran’s I factor, Moran’s I univariate scatter plots were constructed in GeoDa™ for yield, effective depth, clay percentage and N to provide visual indications of the presence of spatial auto-correlation.

3.5.3.2 Breusch-Pagan test

The presence of heteroscedasticity in the error terms was determined by means of the BP test. By running the OLS regression in GeoDa^{TM}, the BP coefficient is produced together with the regression output. The test statistic, which is always non-negative, is derived from the LM test with the null hypothesis of homoscedasticity. Under the null hypothesis, the BP test follows a chi-square distribution with the degrees of freedom equalling the number of parameters (excluding the constant). Large values of test statistics reject the hypothesis that *Y* (yield) is homoscedastic in *X* (effective depth, N, treatment and management zones), with the meaning of "large" depending on the number of variables in *X*. The detection of spatial effects is essential to enable a correct model specification.

3.5.4 Model specification

As the experimental design entails strips of VR and SR application of N, equation 1 below is the most appropriate for analysis to determine the profitability of VR application as a total package. The analysis entails estimation using the data of the three years aggregated together, followed by each year's individual analysis. The following Baseline Model was used in the estimation:

Baseline Model – The regression model to test the technology as a package over time:

$$Y = \alpha_0 + \alpha_1 TRT + \alpha_2 Z_1 + \alpha_3 Z_2 + \alpha_4 Z_3 + \alpha_5 TRT(Z_1) + \alpha_6 TRT(Z_2) + \alpha_7 TRT(Z_3) + \alpha_8 D_2 + \alpha_9 D_3 + \alpha_{10} D_2(TRT) + \alpha_{11} D_3(TRT) + \alpha_{12} Z_1 D_2 + \alpha_{13} Z_1 D_3 + \alpha_{14} Z_2 D_2 + \alpha_{15} Z_2\_D_3 + \alpha_{16} Z_3\_D_2 + \alpha_{17} Z_3\_D_3 + \lambda \ldots \ldots \quad \textbf{(1)}$$

Constraint: $\Sigma \alpha Z_i = 0$

$\boldsymbol{\alpha_1 TRT_1}$ =	1 if variable-rate, or 0 otherwise
$\boldsymbol{\alpha_2 Z_2}$ =	Management Zone 2
$\boldsymbol{\alpha_3 Z_3}$ =	Management Zone 3
$\boldsymbol{\alpha_4 Z_4}$ =	Management Zone 4
$\boldsymbol{\alpha_5 D_2}$ =	Year 2
$\boldsymbol{\alpha_6 D_3}$ =	Year 3
$\boldsymbol{\lambda}$ =	Spatial error

$Z_1 = Z_2 = Z_3 = 0$

$D_1 = D_2 = 0$

The dummy variables are specified in such a way that they sum to 0, i.e. the dropped variable is -1, and the variable in question is 1 or 0. The management zones' dummies sum to zero and the significance of each dummy is interpreted as the difference from the mean value, while the annual dummy variables are interpreted as the difference from the base.

Individual effects of the explanatory variables were assessed through a number of sensitivity tests on the Baseline Model, as reflected by the following equations:

Sensitivity test models:

(i) $Y = \alpha_0 + \alpha_1 TRT + \alpha_2 ED + \alpha_3 ED^2 + \alpha_4 TRT(ED) + \alpha_5 D_2 + \alpha_6 D_3 + \alpha_7 D_2(TRT) + \alpha_8 D_3(TRT) + \alpha_9 D_2(ED) + \alpha_{10} D_3(ED) + \alpha_{11} D_2(ED^2) + \alpha_{12} D_3 ED^2 + \upsilon$ ……...…....(2)

(ii) $Y = \alpha_0 + \alpha_1 N + \alpha_2 N^2 + \alpha_3 Z_1 + \alpha_4 Z_2 + \alpha_5 Z_3 + \alpha_6 NZ_1 + \alpha_7 NZ_2 + \alpha_8 NZ_3 + \alpha_9 N^2 Z_1 + \alpha_{10} N^2 Z_2 + \alpha_{11} N^2 Z_3 + \alpha_{12} D_2 + \alpha_{13} D_3 + \alpha_{14} D_2 N + \alpha_{15} D_3 N + \alpha_{16} D_2 N^2 + \alpha_{17} D_3 N^2 + \alpha_{18} Z_1 D_2 + \alpha_{19} Z_1 D_3 + \alpha_{20} Z_2 D_2 + \alpha_{21} Z_2 D_3 + \alpha_{22} Z_3 D_2 + \alpha_{20} Z_3 D_3 + \upsilon$........................(3)

(iii) $Y = \alpha_0 + \alpha_1 TRT + \alpha_2 N + \alpha_3 N^2 + \alpha_4 Z_1 + \alpha_5 Z_2 + \alpha_6 Z_3 + \alpha_7 NZ_1 + \alpha_8 NZ_2 + \alpha_9 NZ_3 + \alpha_{10} TRT(Z_1) + \alpha_{11} TRT(Z_2) + \alpha_{12} TRT(Z_3) + \alpha_{13} N(TRT) + \alpha_{14} D_2 + \alpha_{15} D_3 + \alpha_{16} D_2 N + \alpha_{17} D_3 N + \alpha_{18} D_2 N^2 + \alpha_{19} D_3 N^2 + \alpha_{20} D_2 TRT + \alpha_{21} D_3 TRT + \alpha_{22} Z_1 D_2 + \alpha_{23} Z_1 D_3 + \alpha_{24} Z_2 D_2 + \alpha_{25} Z_2 D_3 + \alpha_{26} Z_3 D_2 + \alpha_{27} Z_3 D_3 + \upsilon$..(4)

Among the sensitivity test models, the quadratic spatial regression model was viewed essential as it allows for diminishing returns and incorporates spatial effects, spatial auto-correlation and spatial heterogeneity. On the basis of Lau's (1986) criteria for selecting functional forms, the quadratic form was selected for the following reasons:

- The quadratic model has linear parameters, making it easier to estimate using several software packages.
- Quadratic models have the capability to provide second-order estimation to any arbitrary function.
- Quadratic models allow the inclusion of quadratic and interaction terms in the independent variables, which are important in estimating the mean differences.

- Quadratic models do not use many degrees of freedom; however, this is not important in this study due to an adequate number of observations.
- Quadratic models are also well suited to crop biological response, and are widely used in literature.

In the models yield was estimated as a function of the effective soil depth, the applied N and the spatial error. The management zones were also included as dummy variables. For the pooled data, that is, the data of the three years aggregated together, the time dummy variable was used to indicate responses in different years. The variables included differed in different models as formulas indicate.

3.5.5 Data analysis techniques used for VRT in other studies

The spatial component of VRT data necessitates careful consideration of the data analysis technique in order to produce accurate estimates. Different authors and researchers implement various data analysis techniques, each with its merits and shortcomings. The purpose of the analysis – whether estimation, forecasting or interpolation – determines the type of methodology followed. The degree of preciseness also varies from one method to the next. It is vital to take heteroscedasticity and auto-correlation, which often occur in yield data, into account in estimating regression functions in order to provide accurate estimates of parameters for the profitability of precision agriculture. Some of the approaches that have been used in the analysis of precision agriculture data are discussed in the subsections that follow.

3.5.5.1 Spatial econometric technique

The spatial econometric technique was implemented in this study. The main benefit of a precise spatial econometric methodology is that any spatial structure in the data is exploited so that more accurate estimates are obtained for the parameters in a yield response function, and ultimately yield, return and profitability (Anselin, Bongiovanni & Lowenberg-DeBoer, 2004). In yield monitor data, which is the type of data collected for this study, spatial dependence is usually present and can be reflected in regression analysis in two ways: as a *spatial lag* and/or a *spatial error*. OLS estimates prove inefficient, but remain unbiased if spatial error processes are ignored. However, discounting spatial lag processes leads to inconsistent and biased OLS estimates. The incorporation of the spatial weight matrix *w* into the regression model determines the relations between the dependent variable y_i and neighbouring y_i's or error terms, respectively, with regard to lag and error processes (Lambert *et al.*, 2003).

Bongiovanni and Lowenberg-DeBoer (2001) conducted a study aimed at determining the possibility of using spatial regression analysis of yield data to estimate site-specific crop N response and estimating profits for site-specific N management using crop response, as well as comparing profits from site-specific management with uniform rates. The response functions are estimated using spatial econometric techniques, which assume that spatial variability is a relationship among discrete observations. This is undertaken through a three-step procedure. Firstly, the specification tests and diagnostics for the presence of spatial effects are carried out. Secondly, the spatial effects in the econometric model are formally specified, and finally the models that incorporate spatial effects are estimated.

The maize response to N is estimated in a quadratic form for the complete strip, as well as by taking topographic conditions into consideration. This is modelled as:

$$Yield = \alpha_0 + \alpha_1 N_i + \alpha_2 N^2 + \varepsilon$$

where N_i =N rate. Topographic areas are represented as dummy variables.

The presence of spatial dependence in the error model is tested in SpaceStat^{TM}, firstly by Moran's I, which measures spatial auto-correlation in regression residuals. The second diagnostic test for auto-correlation is the Lagrange Multiplier, which tests the slope or gradient of the likelihood function. The last test is the Kelejian-Robinson statistic, which – apart from testing auto-correlation and heteroscedasticity – also indicates the spatial regression model that should be used (error or lag). Diagnostic tests for heteroscedasticity reported by SpaceStat^{TM} include the Breusch-Pegan (BP) and Koenker-Basset (KB) tests, depending on the outcome of the normality test, with the KB test being reported when errors are not normal.

To establish whether maize N response varies according to landscape position, the spatial Chow test is carried out to detect structural instability in spatial regimes. In determining the profitability of VRT N application, spatial regression models of either spatial lag or spatial error are estimated. The estimated coefficients are then used to rank the net returns over N fertiliser and VRT applications costs according to landscape position. Ordinary calculus is used to compute the optimal level of N according to landscape. The returns from VRT are compared to the returns for uniform applications. Returns above fertilizer costs are estimated as follows for the uniform rates and variable rate of N, respectively:

Returns above fertilizer costs ($/ha) $= \Gamma\,(P_C\,[a_i + b_i\,N_i + c_i\,N^2] - P_N N_o)$

$= \Gamma\,(P_C\,[a_i + b_i\,N_i^* + c_i\,N^{*2}] - P_N N_i - F_{VRT})$

P_C = Price of maize (corn)

$_i$ = Landscape area: 1= low; 2= slope E; 3= hilltop; 4= slope W

N_i = Profit maximisation rate of N for the whole field (landscape area for SSM)

P_N = Price of fertilizer, plus interest for six months at 15% annual interest

F_{VRT} = Variable-rate application fee

The focus is primarily on the degree to which the returns for N according to landscape position are higher on average than those of the commonly used uniform rate strategies. The economic analysis to assess whether the added benefits supersede the added variable costs in a typical year, was performed using the partial budgeting tool.

3.5.5.2 Spatial and non-spatial models (OLS and ML)

In the research by Anselin *et al.* (2004), where the objective was, *inter alia*, to estimate the profits for site-specific N management based on the crop responses, both the spatial and non-spatial models are used. Specification tests and the progression thereof are carried out as in Bongiovanni and Lowenberg-DeBoer (2001). Spatial auto-correlation is included in the regression model in two basic ways. In the first instance, auto-correlation is limited to the error term in a regression model, the so-called spatial error model. The second approach relates to spatial auto-correlation in the dependent variable (*y*) in the model itself, the spatial lag model. The SpaceStat™ statistical package was used for all estimation and specification tests.

The maize response to N is estimated as quadratic specification according to landscape position, as:

$$Yield_{ij} = \alpha_i + \beta_i N_{ij} + \beta_i N_{ij}^{\;2} + \varepsilon_{ij}$$

where *i* = landscape; *j* = location within the landscape; *yield* = corn yield and N_{ij} = N rate.

Different models are estimated with the REGIME and AUTO models. A standard REGIME model incorporates landscape positions, and is estimated by OLS. The AUTO model estimated by ML-GHET is a spatial regime model with a spatial auto-regressive error term, as well as group-wise heteroscedasticity.

Various break-even fees are obtained from the economic estimates of the two models. The AUTO model estimated by ML-GHET provides a higher level of confidence in the difference between regions, and therefore more accurate estimates than the REGIME model estimated by OLS. More than double the break-even point of the REGIME model is obtained in the estimates from the AUTO model, suggesting that VRT may be profitable for farmers. This indicates how OLS estimates may be considerably inaccurate when the interaction between yield points is not taken into account, as is done in spatial models (AUTO). Anselin *et al.* (2004) indicate how the two models lead to different economic conclusions. The REGIME model discourages the adoption of VRT N fertiliser, while the AUTO model demonstrates the economic feasibility of VRT. It is therefore clear that the model used can affect economic decision-making regarding precision agriculture.

Voortman and Brouwer (2001) also use the two methods of estimation, namely OLS and ML. In their study for empirical analysis of the simultaneous effects of N, P and potassium (K) in millet production on spatially variable fields, spatial models are estimated with the spatial lag specification, the MLLAG, since it has been determined as the appropriate model based on the specification tests. The coefficients for the spatial lag terms are highly significant in comparison with the OLS coefficients, indicating that OLS results are considerably biased. The importance of taking the spatial processes into account in estimating parameters is brought to the fore by differences in the significance of coefficients when spatial processes are considered, in comparison to the effect of disregarding such processes.

3.5.5.3 REML geo-statistic approach

Lambert, Lowenberg-DeBoer and Bongiovanni (2002) compare maize response to VR N application, estimated on the basis of two regression techniques a geo-statistic approach and a spatial econometric approach, which are then compared to the OLS.

The way in which spatial dependence between observations is defined, distinguishes one approach from another. In the geo-statistic approach, spatial processes are defined by direct representation. Variance-covariance (VC) matrix structures are estimated by stipulating off-diagonal elements as a function of distance that separates all location points. The spatial econometric approach follows an indirect process by defining spatial structure in a VC matrix, whereby the value of a random variable at a location is related to values in neighbouring locations (Lambert *et al.*, 2002).

The REML parameter estimates in the geo-statistic approach are interpreted as generalised least squares estimates adjusted for spatial auto-correlation. REML standard error estimates should, on average, be smaller than the OLS standard error estimates (Lambert *et al.,* 2002).

In the case of the spatial econometric approach, the model is estimated using the SAR ML (error) and SAR (group-wise heteroscedasticity) GHET methods. A comparison between the SAR econometric and the REML geo-statistic regressions indicates that the SAR GHET model has the best fit according to the AIC (Akaike's Information Criterion) and LIK (Log Likelihood) criteria. All topography intercept terms are significant with regard to the OLS, SAR and REML models. The fit of the REML and SAR models is very similar when heteroscedasticity between topographical zones is not taken into consideration. However, the measures of fit in the SAR GHET model improve when heteroscedasticity is taken into account.

By and large, the slope coefficients for all models are alike, with major differences between intercept terms for each topographical region, as well as with regard to the magnitude of parameter significance. The standard errors of the SAR GHET and SAR ML models were 19% less on average than the OLS model. On average, standard errors of the REML group were 3% (spherical) to 7% (exponential) smaller than the OLS standard error estimates. Standard errors of the REML models were 8%, 14% and 20% larger than the SAR GHET and SAR ML standard error for the Gaussian, exponential and spherical specifications, respectively. Since SAR models have a smaller standard error, they have more significant coefficients than either the OLS or REML models (Lambert *et al.*, 2002).

According to Lambert *et al.* (2002), the SAR techniques offer several advantages over the REML method. SAR is a one-step maximum likelihood estimation process, while at least four steps are required in REML. Furthermore, SAR can be implemented based on a small number of observations, while REML requires enough data to estimate the semi-variogram. However, since most soil scientists and agronomists are familiar with geo-statistical concepts, the use of geo-statistic REML may assist in generating confidence in regression analysis (Lambert *et al.*, 2002).

3.6 LIMITATIONS

On-farm research is an indispensable tool in evaluating and validating farming technology under local farming conditions. The best on-station technology rarely performs at the same level under actual farming conditions, due to variability in field-scale soils and essential practical activities at farm level. Researchers' intentions are often to derive generalisable results, which are sometimes in conflict with the farmers' objectives. Farmers' interests lie in improving management for a farming operation, which can translate into increased profits and reduced risk.

Although on-farm comparisons generate such valuable information for decision-making, the results obtained are not statistically valid from the perspective of classical agronomic designs. Such results can be made useful when spatial auto-correlation present in the data is modelled correctly. Employing analysis of variance techniques that do not account for spatial auto-correlations can render the efforts of on-farm comparisons useless. It is also important to note that, due to variability in soils, valid conclusions emanating from different treatments can only be reached with sufficient replications.

Internal validity is essential in experimental design, as it is argued that the results would otherwise become uninterpretable. Although confounding variables have to be controlled to maximise the internal validity of the experimental study, and Leedy (2001) expresses the opinion that definite conclusions about cause-and-effect relationships can only be drawn through a carefully controlled experimental design, full control is impossible in field trials that represent traditional crop production. It is not always possible to control all exogenous variables such as weather conditions and market situations that affect machinery and equipment availability, which, in turn, affect crop cultivation practices and yield. However, measures were taken to maximise internal validity by making sure that only variables of interest were allowed to vary.

Field trials in agriculture are carried out under uncontrollable weather conditions and in an external environment over which the farmers and researchers alike have no control. Furthermore, the ultimate goal of the farmer is to make profit, and management is adjusted along the way to ensure that the farmer achieves his objectives. Other limitations are related to treatment edge effects presented by strip-trial designs, as discussed earlier.

3.7 CONCLUSION

This study replicated maize crop production as it is done on a real farm in order to provide an analysis obtained under the conditions in which farmers operate. The design and the experimental procedures and techniques pursued, allowed investigation into the effects that different application strategies (VR and SR) of inputs have on yield, without causing major disruption to normal farming activities. The design makes it possible for individual farmers to conduct such trials on their farms, but concerns regarding the issue of validity must be kept in mind. It is not possible to control all the extraneous factors in agriculture, since this industry is subjected to the influences of erratic weather conditions and other external factors.

Precision agriculture, also referred to as site-specific management, is indeed site-specific. The responses will vary from farm to farm and even from field to field, and the conclusions can only be made for the field or farm being studied, as no generalisations are possible. The results of this research, as discussed in Chapters 4 and 5, only reflect the situation of the particular field, and no generalisations can be made with regard to the entire farm. This, however, does not imply that this experimental design is unreliable.

Both spatial and non-spatial methods have been used in other studies, with differing parameter estimates. The same consensus was reached with all the spatial models, namely that VR application of nutrients, particularly N, is potentially profitable, while non-spatial models such as OLS estimates were considerably inaccurate, and it was concluded that VRT was unprofitable in these cases. In order to get the best out of the data generated, the best analysis techniques must be selected. Due to the spatial structure present in the data, spatial econometrics had to be used to account for both spatial auto-correlation and spatial heterogeneity. If these spatial effects were not taken into account, inaccurate estimates could be obtained, resulting in imprecise inferences. A correct model that best explains the relationship between the independent variables and a dependent variable, had to be specified. A Baseline regression model with only dummy variables was estimated to test precision agriculture as a package, and various sensitivity analyses were conducted to analyse the Baseline treatment model.

Chapter 4

EXPLORATORY AND DESCRIPTIVE STATISTICS

4.1 INTRODUCTION

This chapter is divided into two sections. The exploratory data analysis for the three years of study is presented first. This is followed by a descriptive statistical analysis of the data. The exploratory data analysis (EDA) employs pie charts and histograms to evaluate the distribution of variables across the field. The relationships between yield and the explanatory variables are evaluated through scatter plots. Moran's I scatter plots are used to assess the presence of any spatial correlation among the variables. Descriptive statistics include the mean, minimum, maximum and standard deviation of yield and soil characteristic variables, the effective depth and the clay percentage. As far as possible, a comparison is made between the two treatments, variable-rate (VR) and single-rate (SR) applications, and the four zones, using primarily the t-tests. This chapter concludes with a short summary of the findings.

4.2 EXPLORATORY DATA ANALYSIS

The aim of EDA is to uncover patterns in the data. EDA is an approach to data analysis that employs a variety of techniques (mostly graphical) to maximize insight into a data set, uncover underlying structure, extract important variables, detect outliers and test the underlying assumptions regarding anomalies (*NIST/SEMATECH, 2004).* Frequency distributions of data variables are indicated by statistical graphs such as pie charts and histograms, while scatter plots demonstrate relationships between the variables.

The EDA proved to be essential, as geo-referenced soil and crop parameters (yield) are the basic data for mapping the variability of yield and soil fertility, and for optimising fertilizer application (Rogasik *et al.*, 1999). The basic data analysis carried out in this chapter entails graphical visualisation of the effective soil depth, clay percentage, yield and nitrogen (N) variables. This graphical visualisation provides insight into the characteristics of the data before any complex and detailed analysis is conducted. For initial examination of spatial variability, yield data were mapped to get a visual indication of large-scale trends in the yield. Yield maps for the VR and SR treatments for the three years of study are attached as Figures 4.1a to 4.1f in Annexure 4.1. The GeoDa™ statistical package was chosen for analysis due to the spatial nature of the data collected, which requires statistical methods specifically designed for geo-referenced data. The Microsoft Excel statistical package was also used for creating pie charts and histograms.

4.2.1 Distribution graphs

Data were analysed and summarised in pie charts and histograms to indicate the distributions across the field. Permanent soil characteristics, effective depth and clay percentage are summarised in pie charts, while yield and N are summarised in histograms. Pie charts graphically organise data to display different values of the permanent soil variables. As the soil variables are permanent, the pie charts represent distributions for the three years of study.

In yield histograms, the observations on the x-axis represent yield points recorded by a yield monitor during harvesting and grouped into different classes, while the frequency of yield observations is indicated on the y-axis. In the case of N, the x-axis corresponds to the application rates during planting, and the y-axis to the number of observations for a particular rate. The advantage of a histogram is that it provides an easy-to-read picture of the location and variation in a data set. However, there are also some limitations associated with histograms that should be borne in mind. Histograms can be manipulated to show different pictures by changing the number of categories to suit the objective. If too few or too many intervals are used, the histogram can be misleading. This is an area that requires some judgement, and perhaps some experimentation, based on the analyst's experience and prior knowledge regarding the variables concerned. The historical yield and input application rates of the previous year were used as a guideline for determining the number of intervals. The intervals for each histogram were determined after experimentation, so that the end result represents the true picture.

4.2.1.1 Effective soil depth distribution

As explained in Chapter 3, the effective soil depth refers to the total thickness of the A- and B-horizons. Percentage distribution of the effective soil depth for the entire field is depicted in the form of a pie chart in Figure 4.1. The distribution remains the same throughout the study years, as this is the permanent soil characteristic. Seven categories were created, ranging from the minimum depth of 50 cm to the maximum depth of 270 cm.

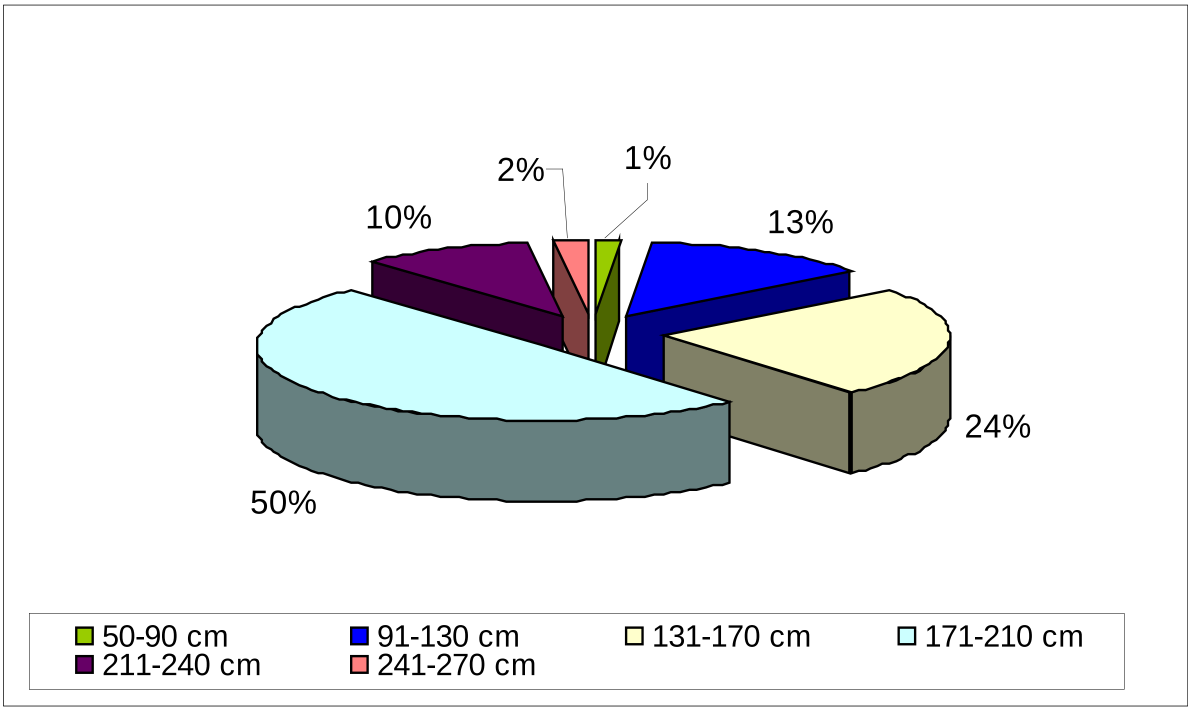

Figure 4.1: Percentage distribution of the effective depth (cm)

As depicted in Figure 4.1, the greater part of the field has an effective depth ranging from 131 cm (24%) to 210 cm (50%). The parts of the field with shallower soils (50–90 cm) are minute, indicating that the majority of soils in the field are above the average recommended depth of 120 cm for good maize production, i.e. yields of approximately 4.5 tons/ha (AGR318, 1998). Very small areas within the field have an effective depth of less than 90 cm (1%) and greater than 270 cm (2%).

4.2.1.2 Distribution of clay percentage

The distribution of clay percentage and the contribution of different types of clay content to the average soil clay content of the field are illustrated in Figure 4.2.

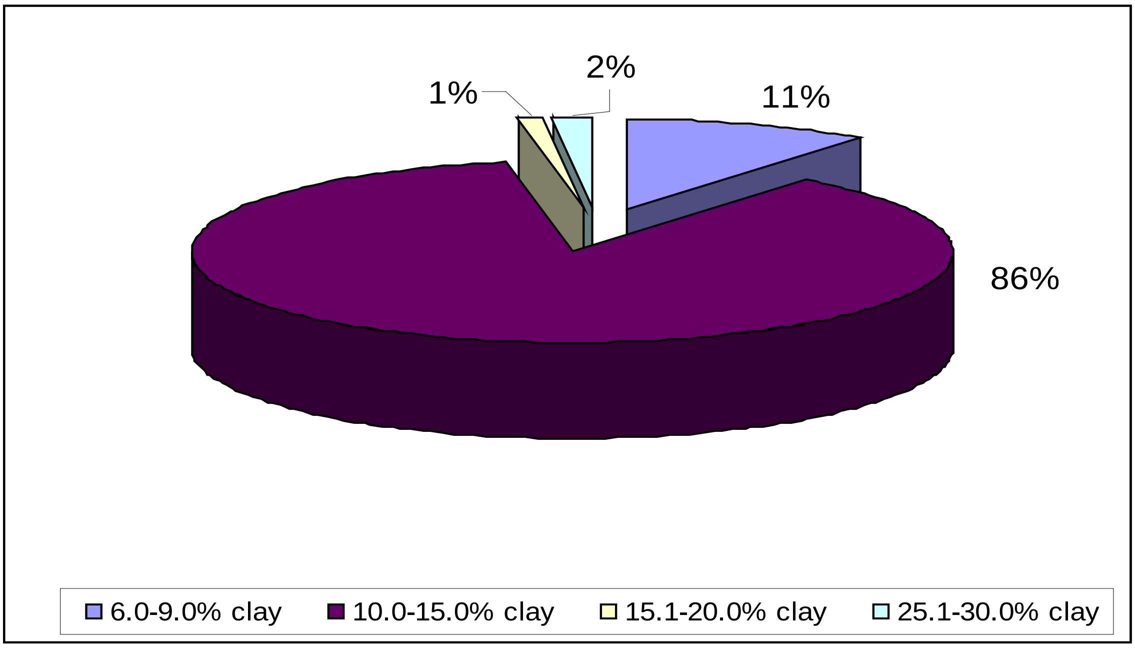

Figure 4.2: Distribution of the clay percentage

According to Figure 4.2, 86% of the soil on the study field has a clay content of 10-15%, while the clay content of 6-9% contributes 11% of the total clay content of the soil. Clay contents of 15.1-20% and 20.1-30%, respectively, contribute 1% and 2% of the total soil clay content. This shows that a very small part of the field has a clay content higher than 20%, indicating that the soils of this field are predominantly of a sandy type, with a clay content of 6–15%.

4.2.1.3 Yield distribution

In the case of the yield, six intervals were chosen, ranging from 1-8 tons/ha and based on historical yield observations. It is essential to mention once again at this stage that the identification of management zones was based on historical yield production, and involved layering the yield maps of the past three years and identifying four yield potential zones. The first zone is the low-potential zone, with a target yield of less than 3 tons/ha. The medium-potential zone, with a potential yield of 3–4 tons/ha, constitutes the second management zone. The third zone is the high-potential zone, with a target yield of between 4 and 5 tons/ha, and finally there is the very high-potential zone, which had a potential yield of more than 5 tons/ha. The yield distribution for the three study years aggregated together, is illustrated in Figure 4.3.

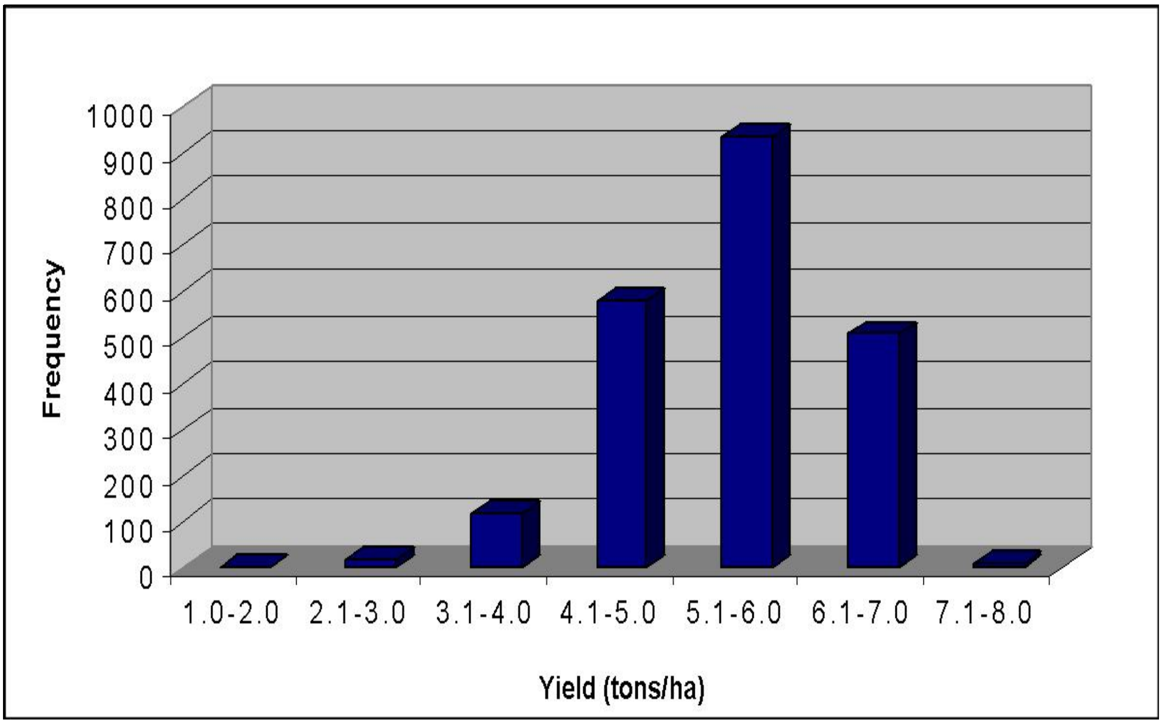

Figure 4.3: Yield distribution for an aggregation of all the study years

Over the three study years, the obtained yield was concentrated in the 4.1-5.0 and 5.1-6.0 tons/ha categories. A very small fraction of yields is categorised as below 3 tons/ha and above 7 tons/ha. A clearer picture can be obtained by looking at yield distributions in each individual year. The yield distribution for year one (2002/2003) is depicted in Figure 4.4.

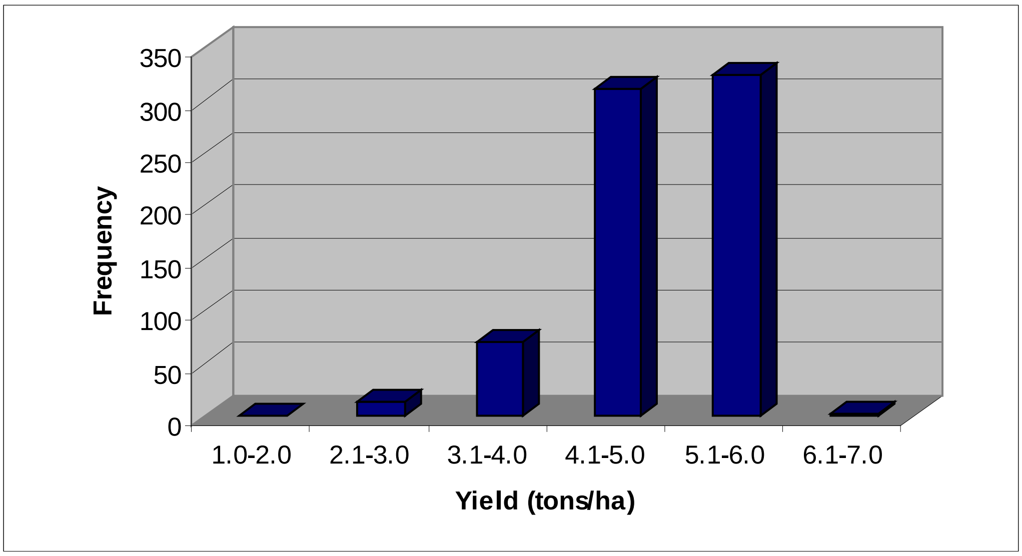

Figure 4.4: Yield distribution (2002/2003)

The fourth and fifth categories of 4.1–5.0 and of 5.1–6.0 tons/ha included a greater number of observations relative to other categories. These are represented by the tallest bars, implying that the majority of observations fell within these categories. In relation to the identification of management zones, these are the high-potential (4-5 tons/ha) and very high-potential (>6 tons/ha) zones. The first and second categories, as well as the last category, which include yields of less than 2 tons/ha and greater than 6 tons/ha (in that order) contained the least number of observations, as shown by the shortest bars. On the whole, the categories of 4.1-5.0 tons/ha and 5.1-6.0 tons/ha (the two longest bars) constituted more than 80% of the yield observations. This shows that the productivity of this field was in the yield range of 4 to 6 tons/ha for the production year 2002/2003. Figure 4.4 shows that during this year, only a small part of the field yielded as low as 1 ton/ha or as high as 7 tons/ha, but the greatest part of the field yielded between 4 and 6 tons/ha.

Yield distributions for the other two years, 2003/2004 and 2004/2005, are presented in Figures 4.5 and 4.6. The same interval was used as in the first year, to allow comparison across the years.

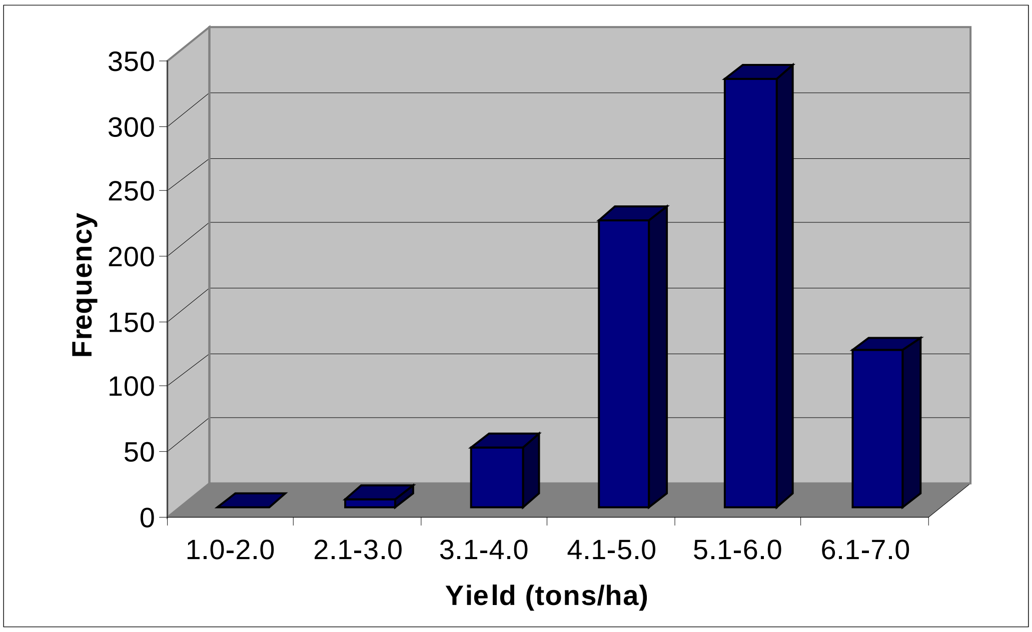

Figure 4.5: Yield distribution (2003/2004)

The same yield trend was observed in the second year as in the first year. About 76% of the yields produced were from the high-potential and very high-potential zones. Yield production in this year ranged mainly from 4–7 tons/ha, with a predominant range of 5.1-6.0 tons/ha. A different picture emerged in the third year, as illustrated in Figure 4.6.

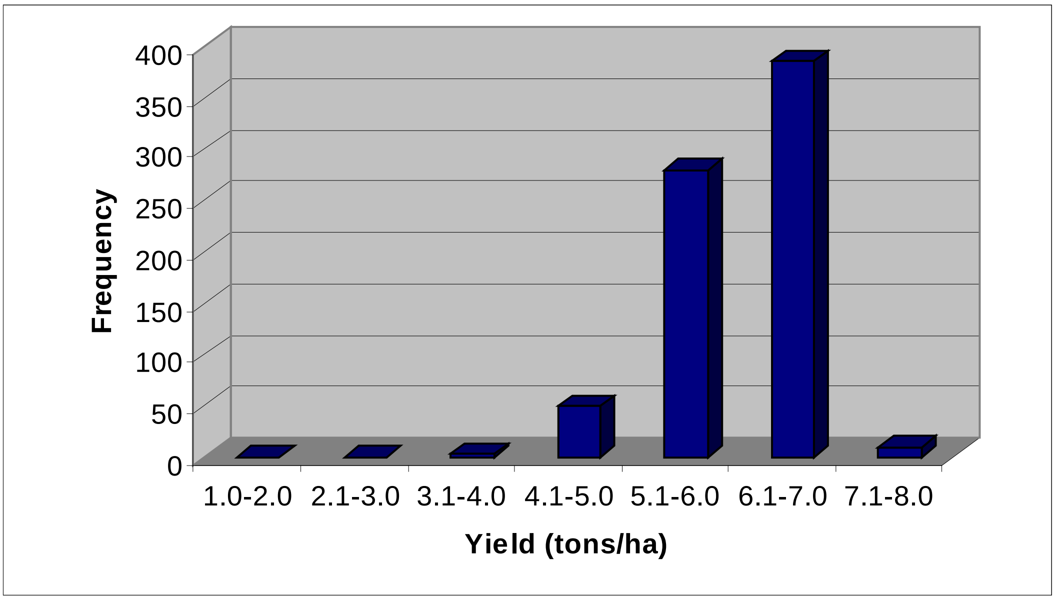

Figure 4.6: Yield distribution (2004/2005)

In the third year, the bulk of the yield was produced predominantly in the very high-potential zone, followed by the high-potential zone, both of which accounted for 91% of the total yield produced. The dominant yield level ranged from 5.1 tons/ha (category 6) to 7 tons/ha (category 7), with the latter category accounting for 53% of the total yield. An increasing yield trend is evident from the first to the third year, with the very high-potential zone gradually increasing in importance with regard to the total yield produced.

4.2.1.4 Applied nitrogen distributions

With regard to the distribution of N application aggregated over a three-year period (Figure 4.7), different application categories are significant. Seven categories were created to represent a wide range of application rates over the three years. The application rates indicated in the histograms are only for the VR treatment. The SR applications are not included, as the rates remained constant.

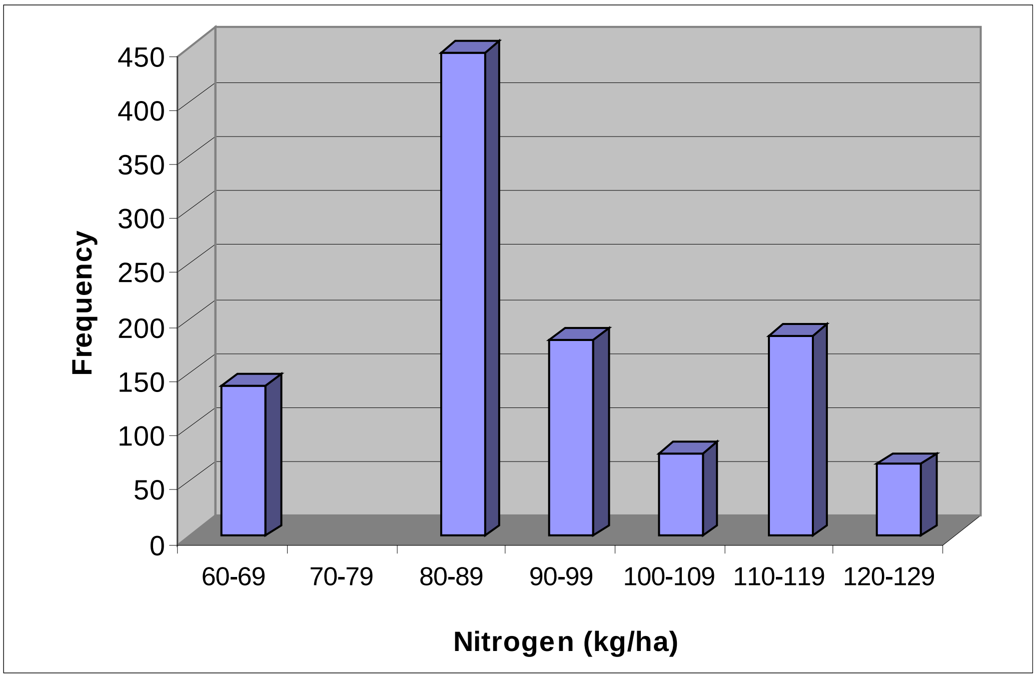

Figure 4.7: Frequency distribution of N application for an aggregation of all study years

Figure 4.7 indicates that the application rates ranging from 80-89 kg/ha were predominant.

The distribution of the applied N in each of the three study years is indicated in Figures 4.8 to 4.10.

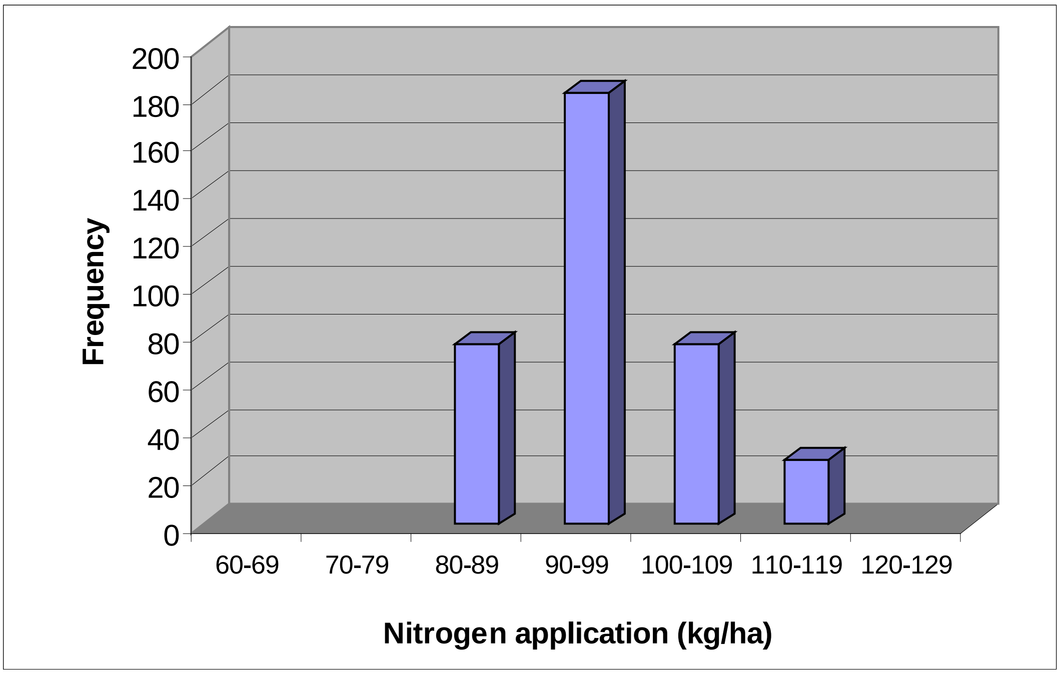

Figure 4.8: Frequency distribution for N application (2002/2003)

Application categories ranging from 90-99 kg/ha represented 50% of the total application rates. These are the rates demarcated for the medium-potential zones. The application rates of 110-119 kg/ha (for very high-potential zones), indicated by the shortest bar, constituted the smallest rate of only 8% of the total rates applied. Rates for the low- and high-potential zones together constituted the remaining 42%. Although no application rates below 70 kg/ha occurred, this rate is included for consistency as lower rates apply in the second year.

The N application frequency distributions for the second and third years are shown in Figures 4.9 and 4.10.

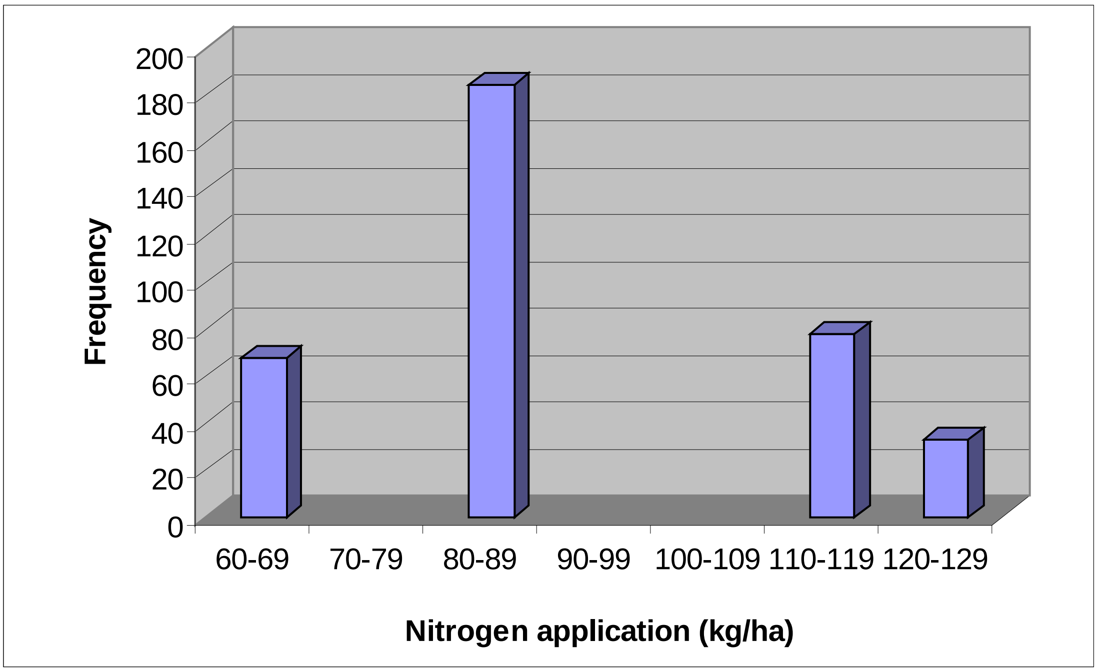

Figure 4.9: Frequency distribution of N application (2003/2004)

In 2003/2004 (year two), N application rates of 80-89 kg/ha were dominant (constituting 51% of the total applications). These are the application rates for the medium-potential zones. The application rates in the very high-potential zone constituted only 9% of the total applications, while the low- and high-potential zones, respectively, constituted 18% and 22%.

Rates applied in 2004/2005 are shown in Figure 4.10.

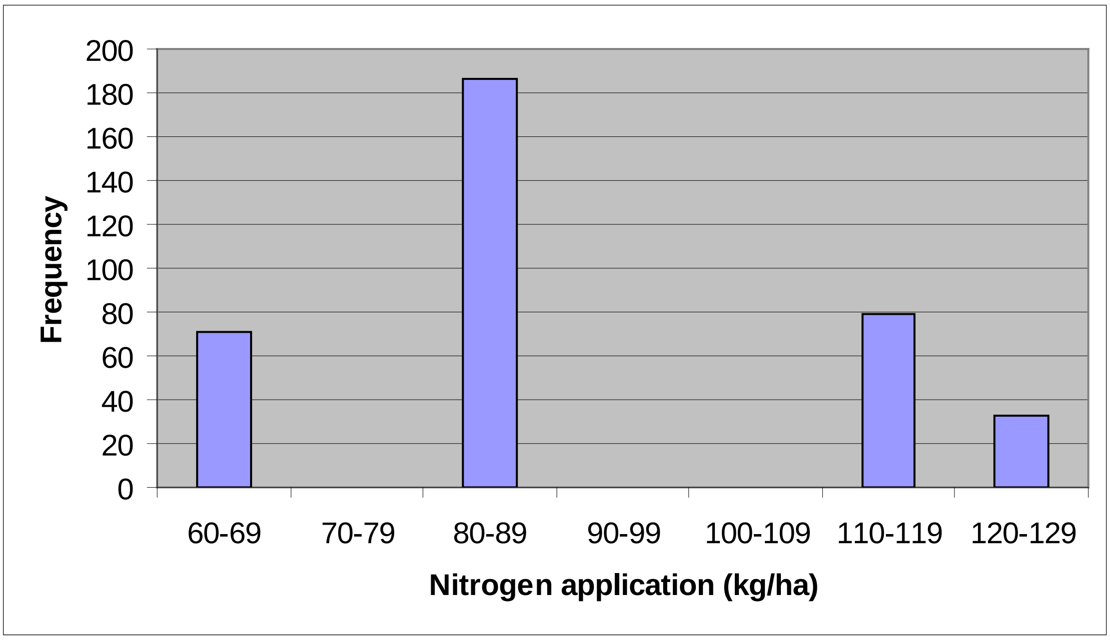

Figure 4. 10: Frequency distribution of N application (2004/2005)

Application rates in the third year ranged from the category of 60-69 kg/ha to the category that encompasses rates of 120-129 kg/ha. The dominant rates were for the medium-potential zones. Over the three study years, the rates that had the highest frequency were for the medium-potential zones, which encompass the greatest part of the field.

4.2.2 Scatter plots

The purpose of scatter plots is to show the type of relationship or correlation that exists between the yield and the independent variables. The scatter plots provide a clue regarding the relationship between the yield and the explanatory variables, as well as the type exhibited. There is a saying in statistics that correlation does not imply causality. In other words, the graph may show that a relationship exists, but it does not and cannot prove that one variable is causing the other. There could be a third factor involved which is causing both, some other systemic cause, or the visible relationship could just be a coincidence (SkyMark, 2004). Therefore, advanced analysis would be required to determine whether any causality exists.

Different scatter plots were created with the yield as the first specified variable on the vertical axis, and the effective soil depth, clay and N as second variables on the horizontal axis. Figures 4.11–13 represent the relationship between yield and the explanatory variables on the basis of the three-year aggregated data. The scatter plots representing relationships in each individual year are attached as Annexure 4.2. The yield is therefore the dependent variable, and the other variables are explanatory variables. GeoDa™ statistical software was used to generate the scatter plots.

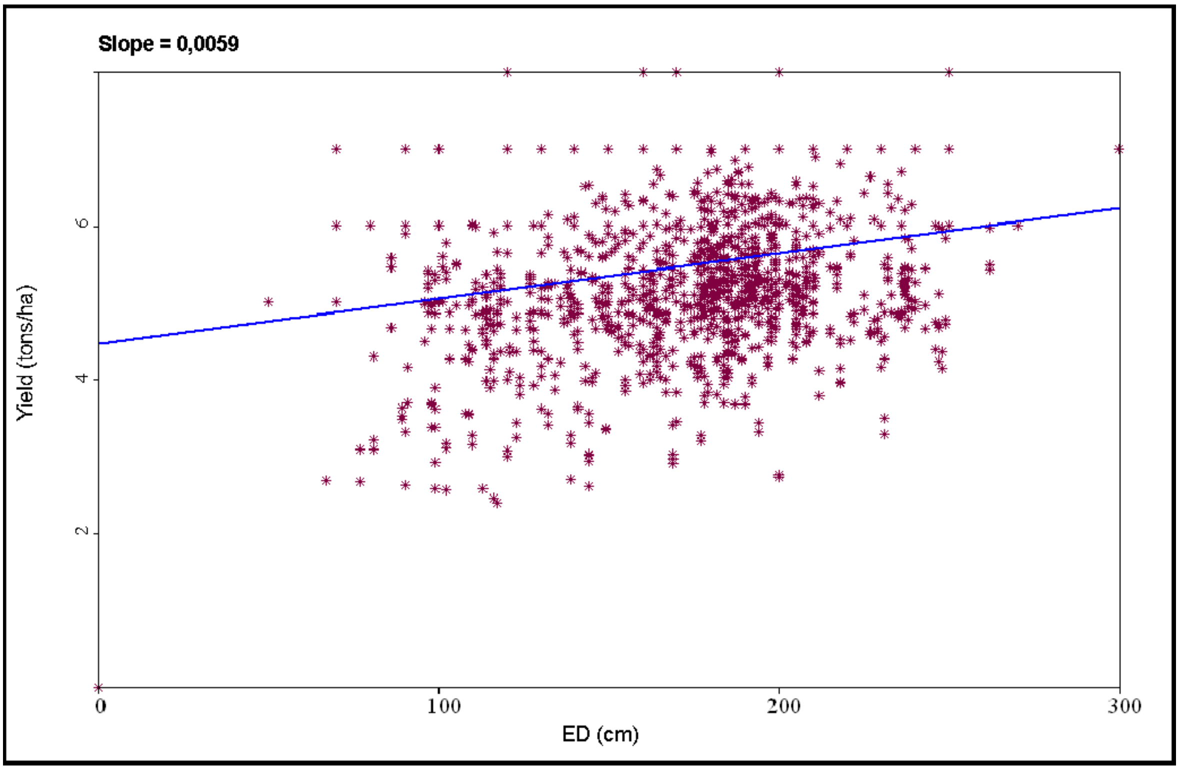

Figure 4.11: Yield-depth scatter plot (aggregated data)

At an effective soil depth of 175-200 cm, a concentration of observations occurs. As shown in the pie chart in Figure 4.1, the depth of the greater part of the field is in this range. The plot indicates a positive relationship between yield and effective soil depth. According to the slope, yield increases by 5.9 kg/ha for each centimetre increase in soil depth. It should be noted that an increase in soil depth can only affect the yield positively up to a point, after which it tapers off to a plateau so that an increase in soil depth does not influence yield anymore.

A positive yield-clay relationship is also evident, as depicted in Figure 4.12.

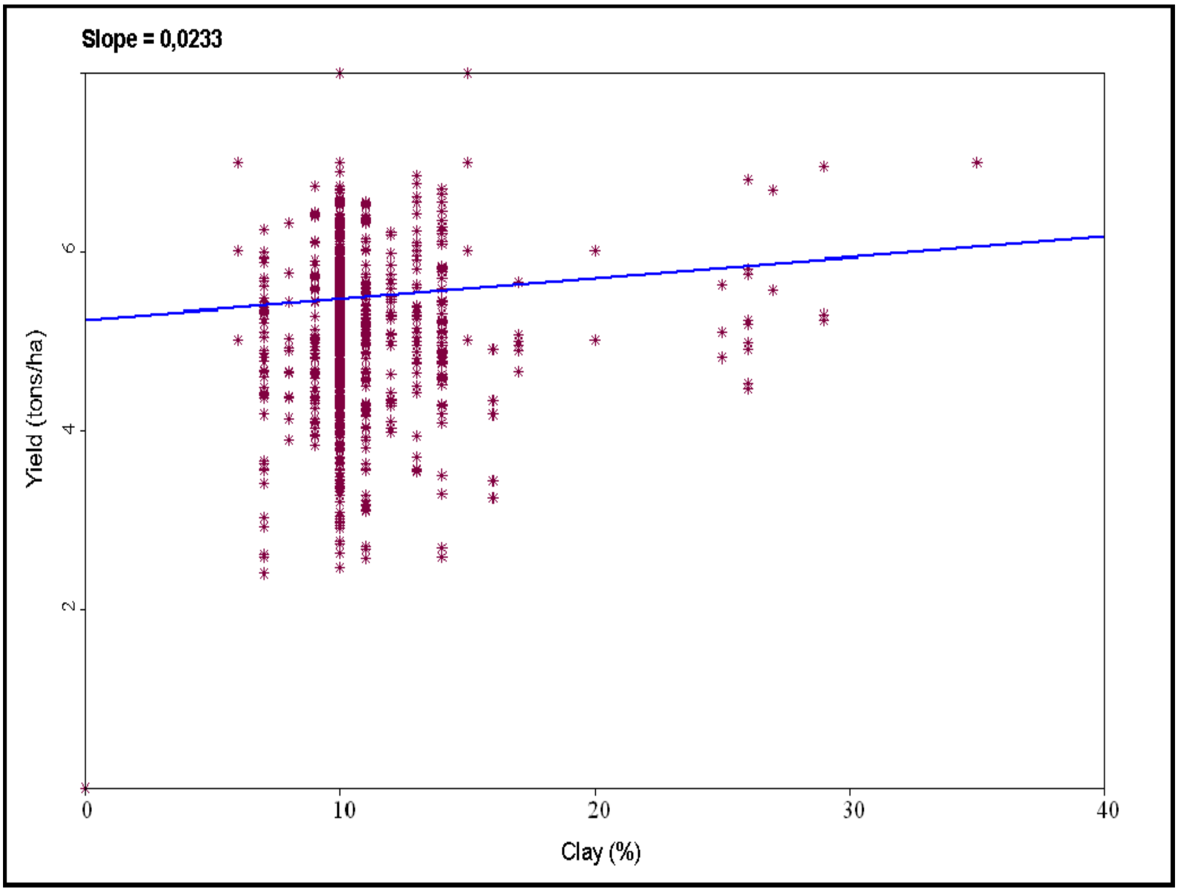

Figure 4.12: Yield-clay scatter plot (aggregated data)

A positive relationship between yield and clay with a slope of 0.0233, means that a 1% increase in clay content leads to an increase in yield by 23.3 kg/ha. Nevertheless, this positive relationship will stop when an excessive clay percentage begins to impact negatively on the yield. The majority of the observations were concentrated at clay percentages of between 6 and 15%, and it can be concluded that yield responses for this field are at their highest at this clay content level. Concentration can be observed at a clay content of 10%, which is the average clay content for the field. The clay content of 6-15% was recorded as clay content available in a greater part of the field during analysis of frequency distribution in Figure 4.2.

A positive correlation existed between yield and the management variable N in the three-year aggregated data, as illustrated in Figure 4.13.

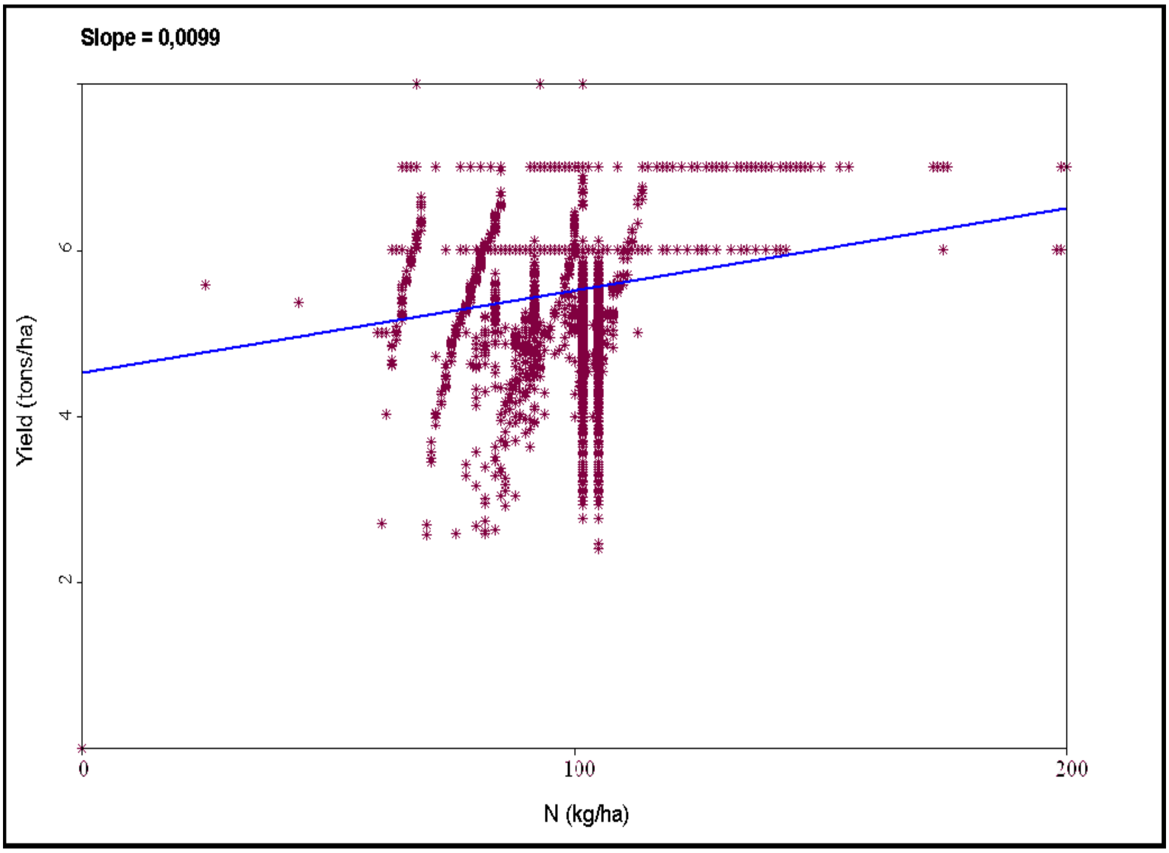

Figure 4.13: Yield-nitrogen scatter plot (aggregated data)

The scatter plot is not as scattered as it should be because distinct N rates were applied, and this creates discrete values. A positive relationship is observed between N and yield, indicating that N has a positive effect on the yield. The slope of 0.0099 implies that, up to a certain point, a kilogram increase in N leads to a yield increase of about 0.01 tons/ha.

In general, positive correlations were observed between yield and the explanatory variables, both permanent and managed, in the pooled data for the three study years. In each of the three years (2002/2003, 2003/2004 as well as 2004/2005) the same positive correlations were observed between yield and these variables. With these kinds of plots, an ordinary least-squares linear regression fit is superimposed on the scatter plot and the corresponding slope listed at the top of the graph. This kind of preliminary analysis does not correct for spatial correlation of the data or heteroscedasticity, but provides some indication of the relationships between the explanatory variables and yield, and positive coefficients are expected in regression analysis. The slope coefficients obtained from the relationship between yield and the explanatory variables in each of the three years are summarised in Table 4.1.

TABLE 4.1: SLOPE COEFFICIENTS OF THE RELATIONSHIP BETWEEN YIELD (KG/HA) AND EXPLANATORY VARIABLES

Variable	2002/2003	2003/2004	2004/2005
Depth (cm)	5.5	9.6	4.8
Clay (%)	19.6	22.7	10.5
N (kg/ha)	9.5	12.9	2.1

The slopes can be interpreted as a preliminary estimate of the amount by which yield in kilograms changes (increases) as each of the variables changes (increases**)** by one unit**.** The interpretations of the slope coefficients are the same as discussed for the aggregated data. The estimated slopes are preliminary in many respects, but particular in the sense that they are derived from univariate regressions. The estimated slopes in Table 4.1 do not allow for the presence of multiple factors.

4.2.3 Moran's I scatter plots

Moran's I, a measure of spatial variability, was employed to determine and measure any spatial correlation in the data. The Moran's I scatter plots provide a visual exploration of spatial autocorrelation; while the Moran's I value measures the degree of spatial association (Anselin, 2002). These scatter plots were produced when conducting a univariate Moran analysis. As described by Anselin (2002), Moran analysis indicates "the spatial lag of the variable on the vertical axis and the original variable on the horizontal axis" (spatial lag being the value of its neighbours). The units on the graph are equivalent to the standard deviations, as the variables are standardised. When values at neighbouring locations are similar, the Local Moran's I statistic will be positive, and negative if they are dissimilar. The weight matrices were created using the Euclidean distance, which ensured that each observation has *at least* one neighbour in each of the years. These are the limits (threshold distances) computed by GeoDa^{TM} to guarantee that each observation has at least one neighbour.

In Univariate Moran's I scatter plots, the Moran's I statistic evaluates spatial autocorrelation in a single variable during or at a single period or point in time. In the four quadrants of the graphs (Figures 4.14-17), categorisation of four types of spatial autocorrelation in yield are depicted. In the case of positive spatial autocorrelation, the quadrants are high-high (upper right) and low-low (lower left). High-low (lower right) and low-high (upper left) quadrants symbolise a negative spatial autocorrelation. Moran's I ranges from -1 to 1, depending on the degree and direction of correlation, with 1 indicating 100% spatial correlation (Anselin, 2003). Spatial autocorrelation on the aggregated data of the three study years is depicted in Figure 4.14, while Figures 4.15, 4.16 and 4.17 depict spatial autocorrelation in yield in each of the three study years.

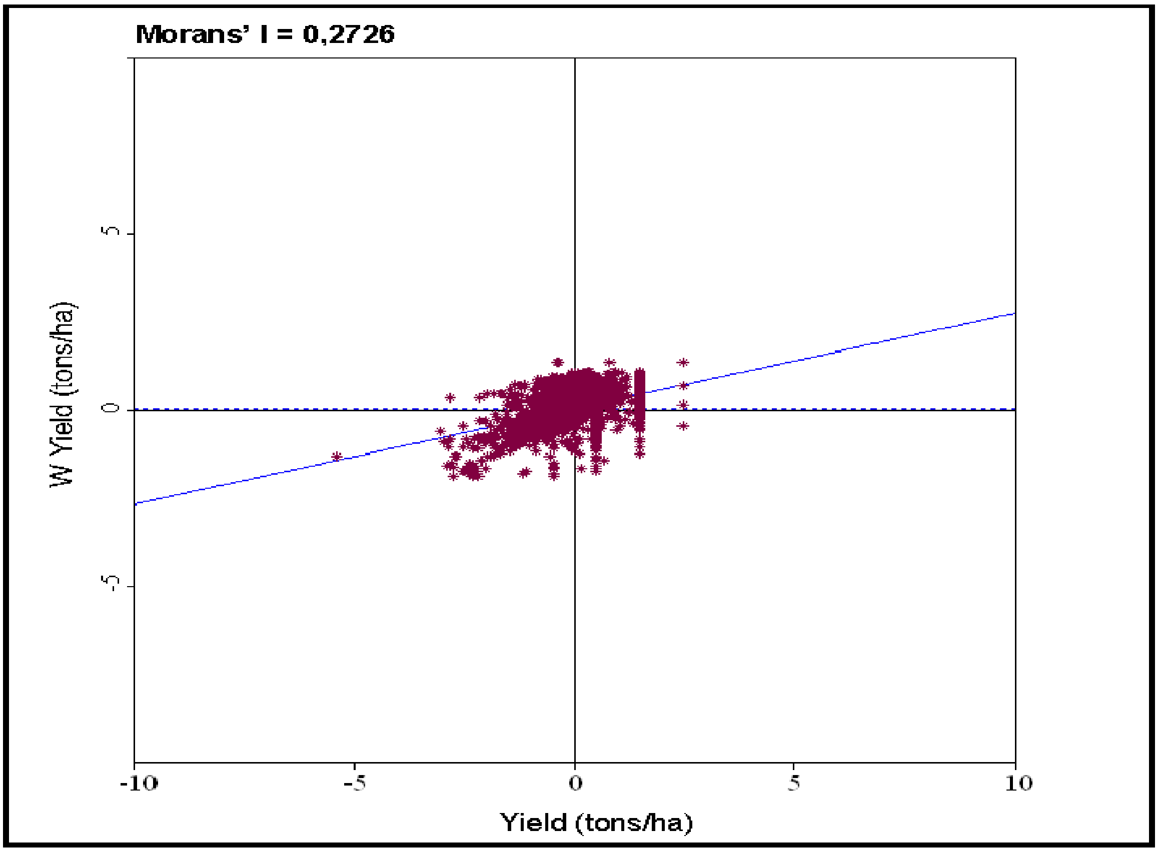

Figure 4.14: Spatial autocorrelation in yield (aggregated data)

Looking at Figure 4.14, the majority of yield observations for the pooled data fall under the two quadrants representing positive spatial autocorrelation, and the trend clearly demonstrates that a positive relationship or autocorrelation exists between neighbouring yield observations. This implies that high yield values have neighbours with high yield values, and this can be attributed to the fact that management zones were allocated in such a way that each management zone is likely to contain observations of similar values.

The Moran's I scatter plot for the first year indicates a high spatial autocorrelation between yield points, as indicated in Figure 4.15.

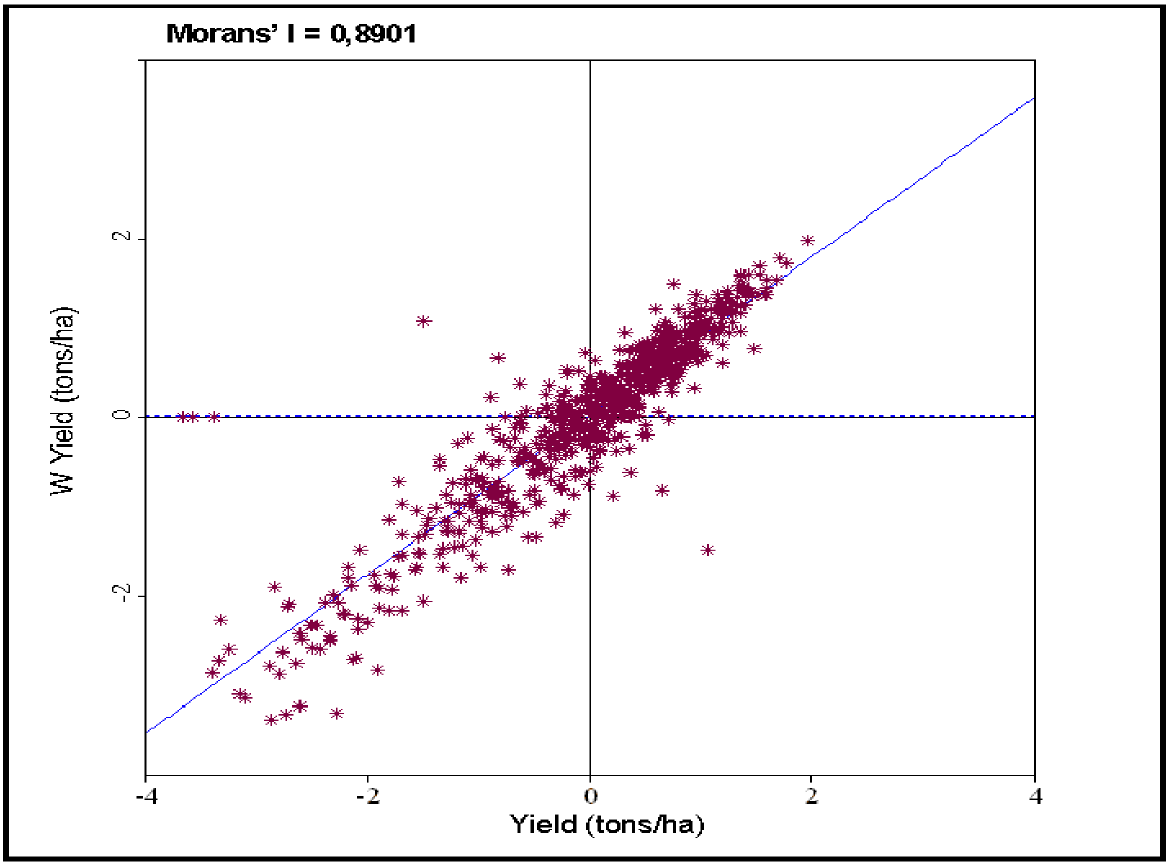

Figure 4.15: Spatial autocorrelation in yield (2002/2003)

Spatial autocorrelation was identified in years two and three as well, as shown in Figures 16 and 17.

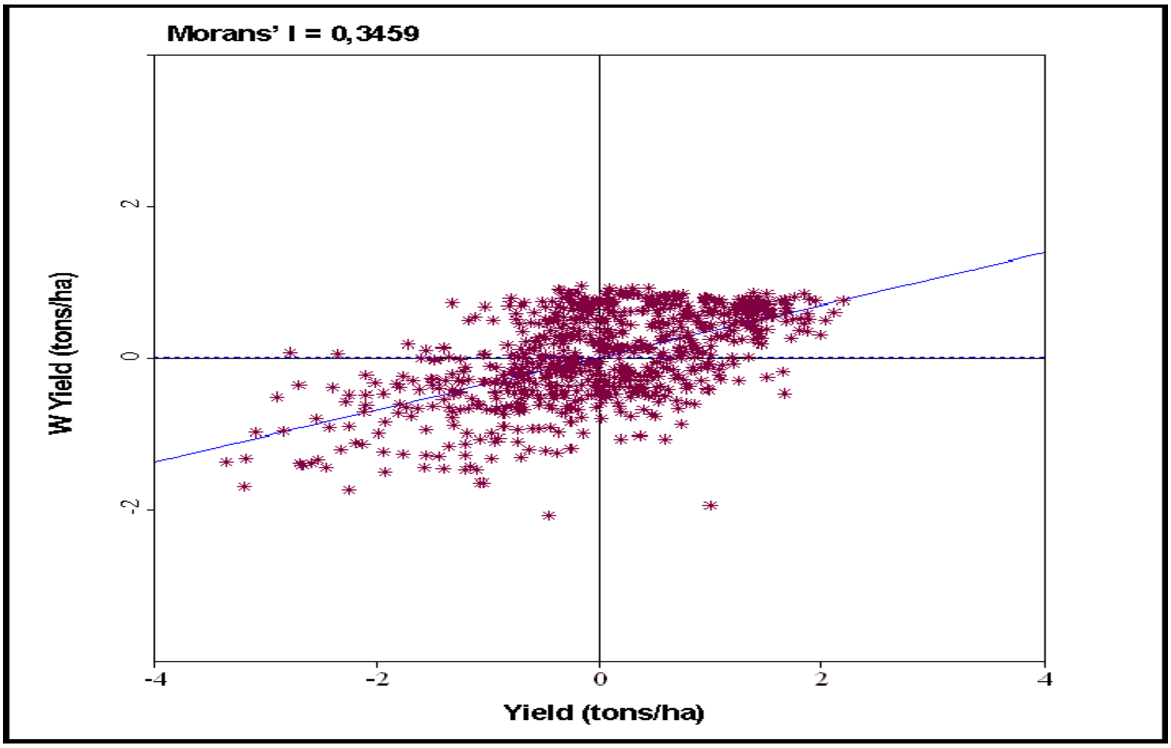

Figure 4.16: Spatial autocorrelation in yield (2003/2004)

A clear positive spatial autocorrelation is still observed in the second-year yield data (Figure 4.16).

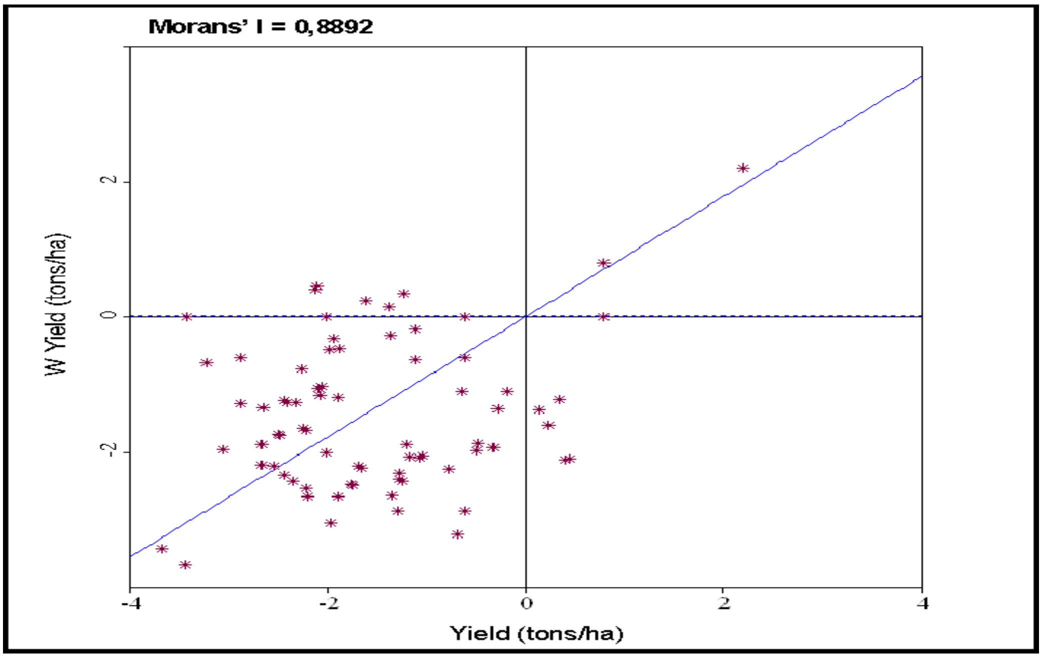

Figure 4.37: Spatial autocorrelation in yield (2004/2005)

Spatial autocorrelation is still positive in the third-year yield data, as depicted in Figure 4.17.

The Moran’s I statistics represented by the slopes of the regression lines, which are listed at the top of the graphs, indicate high and positive spatial autocorrelation of maize yields, with a Moran's I value of 0.2726 calculated for the aggregated data. Moran’s I values are 0.8901, 0.3459 and 0.8892, respectively, for the first, second and third years. As a result, spatial autocorrelation must be taken into account in the regression analysis of data, so that meaningful results can be obtained.

Even though some scatter plots depict outliers of negative spatial autocorrelation, these are cancelled out by the positive relationships since all the Moran’s I statistics calculated are positive, implying the presence of positive spatial autocorrelation. Negative spatial auto correlation implies that low-yielding areas are found within high-yielding areas. With this kind of information, the role of spatial autocorrelation must be taken into account in additional analysis.

4.3 DESCRIPTIVE STATISTICS

The aim of this section is to summarise the data collected in a clear and understandable way by employing the descriptive numerical summaries. Descriptive statistics are used to describe the basic characteristics of the data, with the aim of providing simple summaries and a foundation for inferential analysis. The variables include yield as the dependent variable; the explanatory variables are the effective soil depth, the clay percentage of the A-horizon and N applied in kilograms per hectare. Another explanatory variable is treatment (VR and SR applications of N), represented by dummy variables (1=VR, 0=SR). Descriptive statistics include the mean, minimum, maximum and standard deviation of yield and soil characteristic variables, effective soil depth and the clay percentage. As far as possible, a comparison is made between the two treatments, VR and SR applications, as well as between different management zones. Management zones are different areas within the field, grouped together according to their yield potential, as explained in Chapter 3. Mean differences between different variables are also analysed using the *t*-tests.

4.3.1 Descriptive statistics of yield

Descriptive statistics of yield obtained in this study are presented in Table 4.2. The mean, minimum, maximum and standard deviation of the yield for the entire field are presented, as well as SR and VR strategies for each of the study years. Descriptive statistics of soil characteristic variables are attached as Tables 4.1a, b and c in Annexure 4.3. Only the most important comparisons are discussed in the text.

The average yield obtained over the three-year study period was 5.49 tons/ha. A minimum of 2.39 tons/ha and a maximum of 8.0 tons/ha were computed, while the standard deviation was 1.0. With regard to the two strategies, average yields were 5.71 tons/ha for the SR and 5.53 for VR. Average yields obtained throughout the study years differed from one year to the next. The average yield of the entire field in the first year was 4.8 tons/ha, 5.21 tons/ha in the second year, and 6.44 tons/ha in the third year. The standard deviations for the three years were 0.66, 0.79 and 0.71, in chronological order. In the first year, yield ranged from a minimum of 2.36 tons/ha to a maximum of 6.1 tons/ha. No difference in average yield was observed between the VR and SR application strategies, as the average yields were 4.80 tons/ha for both treatments. Yield variations were observed between different management zones with different treatments. Average yields for the low-, medium-, high- and very high-potential zones with VR application were 4.95, 4.69, 4.84 and 5.02 tons/ha, in that order. In the case of SR application, average yields in corresponding zones were 4.93, 4.73, 4.83 and 4.89 tons/ha, respectively. The range variables for these yields are summarised in Table 4.2.

TABLE 4.2: SUMMARY STATISTICS OF YIELD

2002/2003	*VR Z_1*	*VR Z_2*	*VR Z_3*	*VR Z_4*	*SR Z_1*	*SR Z_2*	*SR Z_3*	*SR Z_4*	*Total SR*	*Total VR*
Mean	4.95	4.69	4.84	5.02	4.93	4.73	4.83	4.89	4.80	4.80
Standard Deviation	0.60	0.78	0.37	0.33	0.62	0.78	0.41	0.43	0.66	0.66
Minimum	2.58	2.57	3.91	4.35	2.39	2.45	3.69	3.82	2.39	2.57
Maximum	5.71	6.10	5.44	5.57	5.77	6.10	5.70	5.65	6.10	6.10
2003/2004	*VR Z_1*	*VR Z_2*	*VR Z_3*	*VR Z_4*	*SR Z_1*	*SR Z_2*	*SR Z_3*	*SR Z_4*	*Total SR*	*Total VR*
Mean	5.66	5.46	5.88	5.66	4.96	4.75	4.82	4.88	4.82	5.61
Standard Deviation	0.69	0.79	0.58	0.68	0.53	0.76	0.43	0.42	0.63	0.74
Minimum	2.69	2.55	4.29	3.99	3.16	2.76	3.69	3.82	2.76	2.56
Maximum	6.64	6.95	6.89	6.76	5.77	6.10	5.70	5.65	6.10	6.95
2004/2005	*VR Z_1*	*VR Z_2*	*VR Z_3*	*VR Z_4*	*SR Z_1*	*SR Z_2*	*SR Z_3*	*SR Z_4*	*Total SR*	*Total VR*
Mean	6.58	6.42	6.87	5.66	6.54	6.41	6.87	4.88	6.40	6.48
Standard Deviation	0.65	0.58	0.43	0.68	0.71	0.58	0.43	0.42	0.77	0.65
Minimum	5.00	5.00	6.00	3.99	4.00	5.00	6.00	3.82	3.82	3.99
Maximum	8.00	8.00	8.00	6.76	8.00	8.00	8.00	5.65	8.00	8.00

VR- Variable-rate treatment
SR- Single-rate treatment
Z_1 – Zone 1 (Low-potential zone)
Z_2 – Zone 2 (Medium-potential zone)
Z_3 – Zone 3 (High-potential zone)
Z_4 – Zone 4 (Very high-potential zone)

Yields in the second year ranged from a minimum of 2.56 tons/ha to a maximum of 6.95 tons/ha. Average yields from the VR application treatment were 0.79 tons/ha more than average yields under the SR. Average yields were 4.82 tons/ha for SR and 5.61 tons/ha for the VR. With regard to average yields in different management zones, VR treatments outperformed all the SR treatments in all management zones. Consequently, the average yields for the VR treatment were 0.79 tons/ha more. Descriptive statistics for yields in different zones under the two treatments are presented in Table 4.2. The average yield obtained for the entire field was 5.21 tons/ha in the second year – 0.41 tons higher than the average for the first year.

In year three, the average yields between the treatments were slightly different, with VR performing better than SR in Zone 4 only. VR outperformed SR by 0.78 tons/ha in Zone 4. In the other three zones, the differences were small. The yields for the whole field averaged 6.44 tons/ha. A minimum of 3.82 tons/ha and a maximum of 8 tons/ha yields were obtained. On the whole, average yields for the three study years were highest in the third year, with a yield increase trend observed from the first to the third year. It is presumed that the effects of on-farm comparisons would also contribute to yield changes in different years, since management skill and knowledge increased with the years, leading to better decisions and an improvement in yields. The yield trends over the three-year study period for VR and SR treatments are illustrated in Figures 4.18 and 4.19.

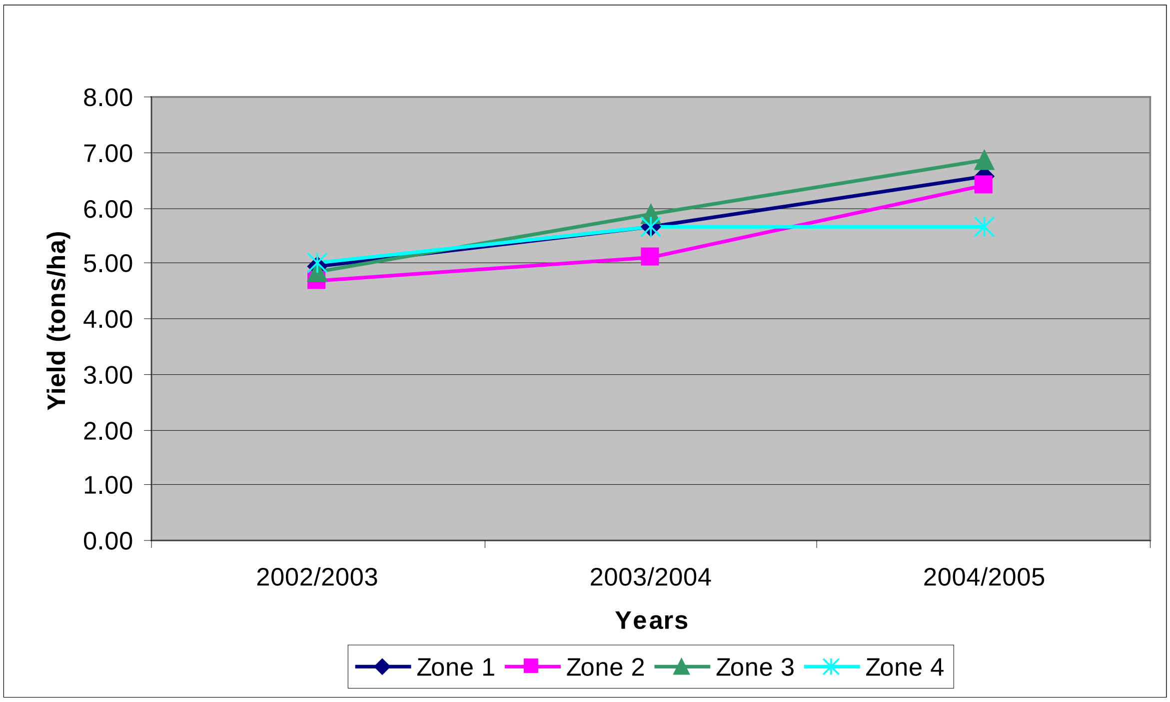

Figure 4.18: VR yield trend

Figure 4.18 indicates that there has been an increasing trend in yields obtained from the VR treatment. The very high-potential zone (Zone 4) did not exhibit a recognizable increase in yields over the three study years, and stabilised to a certain extent from the second to the third year. The same increasing trend also occurred with regard to the SR treatment, as depicted in Figure 4.19.

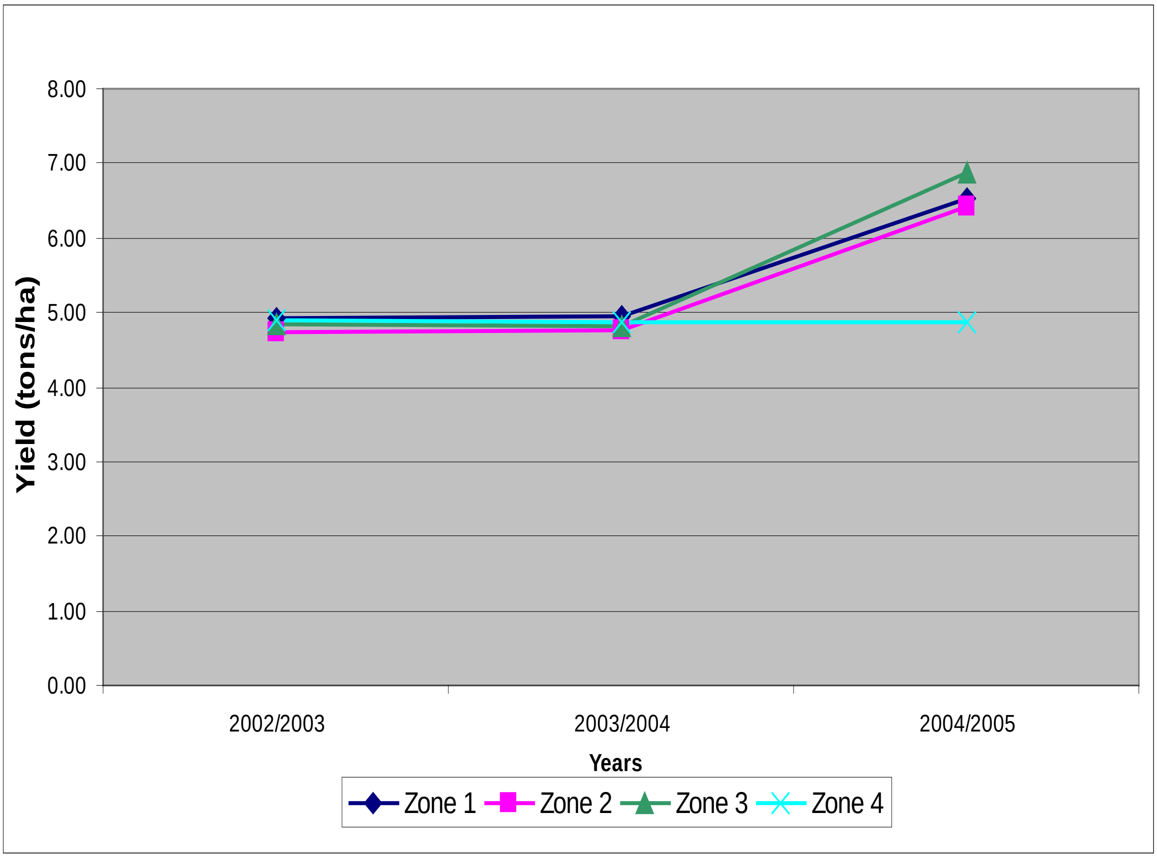

Figure 4.19: SR yield trend

In contrast to the VR treatment, Figure 4.19 indicates that, in the case of the SR treatment, all the zones did not display an increase in yield from the first year to the second year, increasing substantially from the second to the third year (with the exception of Zone 4, which stabilised in the course of the study years). The high-potential zone performed the best in all the years, while the very high-potential zone did not produce according to expectation.

The standard deviation is another important statistic describing the behaviour of the data set. It indicates the average amount by which yield observations for the entire field, treatment (VR/SR) or zone differ from the mean, ignoring the sign of difference. The standard deviations of yield in different years, treatments and management zones are included in Table 4.2. No clear indication was observed regarding the yield variability in different zones when considering the standard deviation.

4.3.2 Descriptive statistics of soil properties

Permanent soil properties are important, as they are a given that the farmer cannot manipulate; he has to plan his management to accommodate them. Furthermore, permanent properties remain constant for many years. Descriptive statistics for the soil properties vary modestly throughout the field. The average clay content of the field is 10.62% and the minimum is 6%, while the maximum is 35%. With regard to the effective soil depth, the mean effective soil depth varies in different management zones. The high-potential zone has the highest mean effective soil depth of 197 cm, followed by the very high-potential zone at 178 cm and the low-potential zone at 173 cm. The effective depth may have contributed to the better performance of the high-potential zones relative to the other zones. The least effective soil depth occurs in the medium-potential zone, with an average effective depth of 164 cm. The medium-potential zone also happened to be the lowest-producing zone. The average effective soil depth for the whole field is 174 cm. It may be concluded that, in demarcating management zones, permanent soil properties such as the effective depth should play an important role as they directly affect the performance of the zones. Summary statistics of soil characteristics are provided in Tables 41a, b and c in Annexure 4.3.

4.3.3 Descriptive statistics of inputs

The relationship between the observed yields and the applied inputs had to be investigated. Based on the aggregated data, the average N applied was 97 kg/ha for VR, and 103 kg/ha for the SR. The standard deviation was high in the VR treatment (19.57 kg/ha), compared to the standard deviation of only 1.73 in the SR treatment. This was expected, as the SR is a single-rate application treatment. In the first year, a constant amount of 105 kg/ha of N was applied in the SR treatment. This is equal to the minimum and maximum amounts, since a constant amount was applied on this treatment. The average N applied was 96 kg/ha under the VR treatment, while the minimum and maximum were 84 kg/ha and 110 kg/ha, respectively. An average of 96 kg/ha is consistent with the exploratory data analysis, where the majority of the observations fell under the 90-99 kg/ha application rates.

In the second and third years, 102 kg/ha was applied as a constant amount. The average N applied for the VR treatment was 97 kg/ha in the second year, while it was 99 kg/ha in the third year. The minimum of 69 kg/ha was applied in the second and third years, while the maximum amounts were 123 kg/ha and 127 kg/ha, respectively. The philosophy of on-farm trials is evident here, as the results obtained in the preceding years are used to adapt management practices as time goes on.

4.3.4 Mean differences between variables: the t-tests

The *t*-test is a statistical significance test used to compare differences between the means. *T*-tests were used to ascertain whether there are any statistical differences between the means of the permanent soil properties (clay percentage and effective depth) in different zones and between the two treatments, as well as within the three years of study. Statistical differences between the mean yields in different zones, among the treatments as well as between the years, were also established.

Statistical differences between the means of the variables were analysed, based on 5% and 10% levels of significance. To determine the statistical differences in the average clay percentage between different zones, *t*-tests for sample means that assume either equal or unequal variances were used, depending on the results of the *F*-test pertaining to variances between the samples. The mean clay percentages between different zones are not statistically different in the three years of study. The mean clay percentage is also not statistically different between the two treatments in all the three years. The results of the mean differences between clay contents in different management zones using the *t*-tests, are attached as Table 4.2 in Annexure 4.4.

With regard to the effective soil depth, there are statistical differences in the mean effective depths between zones for all three study years. Statistical differences between mean effective depths of different management zones were found in all but Zones 1 and 4 in each of the three years, even though Zone 4 has a higher arithmetic average effective depth than Zone 1. There are differences in average effective depth between Zones 1 and 2; Zones 1 and 3; Zones 2 and 3, as well as between Zones 3 and 4. The effective depth in Zone 3 is higher than in the other two zones. The statistical differences between the mean effective depths of the management zones are shown in Table 4.3 of Annexure 4.4.

With regard to N, the average N applied in the management zones is statistically different at 1% level in all the three years, as this is a pre-determined management variable. As far as yields are concerned, the statistical differences between different zones with regard to the mean yields in the three years are summarised in Table 4.3.

TABLE 4.3: MEAN YIELD DIFFERENCES

Zone comparisons	Means	Statistically difference	Best performer
Year 1 – 2002/2003			
Z_1:Z_2	4.94: 4.71	Yes	Z_1
Z_1:Z_3	4.94: 4.83	No	Z_1* (mathematically)
Z_1:Z_4	4.94: 4.95	No	Z_4* (mathematically)
Z_2:Z_3	4.71: 4.83	Yes	Z_3
Z_2:Z_4	4.71: 4.95	Yes	Z_4
Z_3:Z_4	4.83: 4.95	Yes	Z_4
Year 2 – 2003/2004			
Z_1:Z_2	5.31: 5.11	Yes	Z_1
Z_1:Z_3	5.31: 5.34	No	Z_3* (mathematically)
Z_1:Z_4	5.31: 5.27	No	Z_4* (mathematically)
Z_2:Z_3	5.11: 5.34	Yes	Z_3
Z_2:Z_4	5.11: 5.27	Yes	Z_4
Z_3:Z_4	5.34: 5.27	No	Z_3* (mathematically)
Year 3 – 2004/2005			
Z_1:Z_2	6.56: 6.42	Yes	Z_1
Z_1:Z_3	6.56: 6.87	Yes	Z_3
Z_1:Z_4	6.56: 5.27	Yes	Z_1
Z_2:Z_3	6.42: 6.87	Yes	Z_3
Z_2:Z_4	6.42: 5.27	Yes	Z_2
Z_3:Z_4	6.87: 5.27	Yes	Z_3

*The zone may be mathematically higher than, but statistically the same as the paired counterpart.

In year one, average yields differed statistically between Zone 1 and 2; Zone 2 and 3, as well as between Zone 2 and4, with Zones 1 and 3 producing higher yields than their paired counterparts. However, no statistical differences were found between the average yields of Zones 1 and 3; as well as Zones 1 and 4.

In year two, average yields differed statistically between Zones 1 and 2, Zones 2 and 3 and between Zones 2 and 4. The differences in yields were not found in Zones 1 and 3, Zone 1 and 4, and Zone 3 and 4. On the contrary, average yields are statistically different between all the zones in year three. Zone 3 performed better than all the other zones in the pairs. Between Zone 1 and 2, Zone 1 outperformed Zone 2; in a pair of Zone 1 and 4, Zone 1 performed better; for the pair Zone 2 and 4, Zone 2 produced higher yields than Zone 4.

In terms of mean yield differences between the two treatments in the first year, yields were statistically the same for all the zones. In the second year, yields under the VR treatment were statistically higher than yields obtained under the SR treatment, all in the same zone. The VR treatment performed better than the SR in all the zones. In the third year, there was a significant difference in the means of the two treatments in Zone 4 only. Statistical differences were found between all the paired zones in year three.

In comparing the mean yields between treatments over the years (Table 4.5 of Annexure 4.4), i.e. the same treatment in the same zones in different years, average yields were not significantly different in a few zones. Yields for the SR treatment of Zone 1 were not different in all three study years; this also applies to the mean yields of Zone 4 between years one and three with regard to SR treatment. In the rest of the paired zones for different years, as shown in Table 4.5 of Annexure 4.4, the mean yields were statistically different at a 1% level of significance.

4.4 CONCLUSION

Permanent soil properties evaluated include the effective soil depth and the clay percentage in the A-horizon. Frequency distributions and descriptive statistics of these soil properties, together with N as an applied input, were determined in preparation for further analysis of their relationship with yield.

The greater part of the field has an effective soil depth ranging from 131-210 cm, and the clay content in 85% of the soil is 10-15%. The frequency distribution and descriptive statistics of the applied inputs vary over the three years, as management was adjusted based on the results of the previous years. The impact thereof can be observed in obtained yields. In the first year, 4-4.9 tons/ha and 5-5.9 tons/ha yield ranges constituted more than 80% of the yield observation, while in the second year, yield production ranged from 4–7 tons/ha, with the 6.0-7.0 tons/ha range dominating. The dominant yield level ranged from

6–8 tons/ha, the former accounting for 46% of the yield in the third year. An increasing yield trend is evident from the first to the third year.

Average yields were the same for the two treatments in year one. When yields per management zone were analysed, the VR treatment performed better for the very high-potential zone than the SR in year one, but this changed in the second year. The VR average yields were higher than the SR, and the VR management zones outperformed the management zones in the SR treatment. However, in the third year, the VR performed slightly better than the SR in Zone 4 only. The *t*-tests indicate that yields vary significantly between different zones, with Zone 2 differing statistically from all the zones and producing the least in all the study years.

A positive relationship was observed in scatter plots between yield and all the descriptive variables. However, spatial correlation was identified in the data, calling for appropriate measures to account for this phenomenon.

Chapter 5

STATISTICAL AND PROFITABILITY ANALYSIS

5.1 INTRODUCTION

Statistical analysis using regression models, as well as profitability analysis using partial budgets, are the techniques used in this chapter to analyse the data collected on the variable-rate (VR) application of nitrogen (N), in comparison with single-rate (SR) application. The results obtained using the traditional statistical analysis method, Ordinary Least Squares (OLS) and the spatial analysis method, Spatial Error (SER) model, are presented and compared. The SER model uses the Maximum Likelihood (ML) estimation. In the presence of spatial variability, namely spatial autocorrelation and spatial heterogeneity, OLS and other traditional analysis methods are unreliable as the assumptions of normality, independence in observations and identically and independently distributed errors, are violated (Lambert, Lowenberg-DeBoer & Bongiovanni, 2004). Spatial regression analysis makes it possible to overcome these limitations, since it takes into account spatial autocorrelation, which is very important in spatial data such as yield data. Profitability analysis entails the methodology used to determine profitability differences between the two application strategies: VR and SR.

5.2 STATISTICAL ANALYSIS

The aim of this statistical or regression analysis is to investigate the relationship between yield and the effective soil depth and N as explanatory variables. The effects of the two application strategies (VR and SR) on yield are also investigated. The effects of the two application strategies, VR and SR, are captured in a treatment (TRT) dummy variable, which assigns a value of 1 to VR and 0 to SR. Besides N and the effective soil depth as explanatory variables, dummy variables for different management zones are also included in a regression model. The aim is to determine whether yields vary spatially, and this spatial variability is captured by different management zones. Zone 1 (Z_1) is the low-potential zone with a yield potential of less than 3 tons/ha, Zone 2 (Z_2) represents the medium-potential zone with a target yield of 3-4 tons/ha and Zone 3 (Z_3) is the high-potential zone with a target yield of between 4 and 5 tons/ha, while Zone 4 (Z_4) is the very high-potential zone with a potential yield of more than 5 tons/ha, and is used as a base variable. The year dummy variable is also included to test the stability of N response for maize over the years. The dummy variable for the zones is specified as differences from the mean, with the constraint that all the dummy variable coefficients must sum to zero, while the year dummy variable is defined in terms of the reference year. Descriptive statistics of variables can play an important role in explaining the regression output, and a thorough discussion in this regard is presented in Chapter 3. To have an overview of the data, the mean values of the different variables are summarised in Table 5.1.

TABLE 5.1: MEANS OF VARIABLES

SR	Yield (ton/ha)	Clay (%)	ED (cm)	N (kg/ha)	*VR*	Yield (tons/ha)	Clay (%)	ED (cm)	N (kg/ha)
Average	4.80	10.60	174.25	105.00	*Average*	4.80	10.62	174.10	96.00
Zone 1	4.93	10.60	175.04	105.00	*Zone 1*	4.95	11.04	171.03	84.00
Zone 2	4.73	10.46	163.41	105.00	*Zone 2*	4.69	10.29	164.73	92.00
Zone 3	4.83	10.83	198.05	105.00	*ZONE 3*	4.84	10.91	196.25	100.00
Zone 4	4.89	10.81	178.11	105.00	*ZONE 4*	5.02	10.89	177.78	110.00
Average	4.82	10.61	174.42	102.00	*Average*	5.61	10.63	174.17	97.00
Zone 1	4.96	10.72	174.39	102.00	*Zone 1*	5.66	10.73	174.32	69.00
Zone 2	4.75	10.43	164.47	102.00	*Zone 2*	5.11	10.45	164.56	85.00
Zone 3	4.82	10.84	194.01	102.00	*Zone 3*	5.88	10.86	193.48	110.00
Zone 4	4.88	10.82	181.55	102.00	*Zone 4*	5.66	10.82	181.55	123.00
Average	6.40	10.66	173.43	102.00	*Average*	6.48	10.65	173.58	99.00
Zone 1	6.54	10.69	173.47	102.00	*Zone 1*	6.58	10.70	174.23	69.00
Zone 2	6.41	10.47	162.34	102.00	*Zone 2*	6.42	10.47	163.01	85.00
Zone 3	6.87	10.97	194.56	102.00	*Zone 3*	6.87	10.97	194.56	114.00
Zone 4	4.88	10.82	181.55	102.00	*Zone 4*	5.66	10.82	181.55	127.00

Even though the standard OLS regression analysis is conducted, the spatial econometric analysis is of paramount importance in this study. Because of the nature of the data collected, which is inherently spatially correlated, it is essential to use the methodology that takes the effects of spatial autocorrelation into account. In yield monitor data, spatial autocorrelation arises from the coincidence of similarity in yield values and location between yield points. A series of steps is conducted in the analysis. Firstly, the diagnostic tests on normality, spatial autocorrelation and heteroscedasticity are carried out. The Baseline Model (Treatment Model), which analyses precision agriculture as a package by assessing the statistical significance of the estimated coefficients, is employed. Three Sensitivity Test models of the Baseline Model are also presented with a view to determining consistency. Regression analysis – using both the OLS and the SER models of an aggregation of the three years' data into a single regression – is presented in the first place, followed by individual analyses for each of the three years of study. The GeoDa™ statistical package for spatial data analysis was used for all estimations and specifications.

The data of the three years aggregated together was estimated with the following formula:

$$Y = \alpha_0 + \alpha_1 TRT + \alpha_2 Z_1 + \alpha_3 Z_2 + \alpha_4 Z_3 + \alpha_5 TRT\,(Z_1) + \alpha_6 TRT\,(Z_2) + \alpha_7 TRT\,(Z_3) + \alpha_8 D_2 + \alpha_9 D_3 + \alpha_{10} D_2\,(TRT) + \alpha_{11} D_3\,(TRT) + \alpha_{12} Z_1 D_2 + \alpha_{13} Z_1 D_3 + \alpha_{14} Z_2 D_2 + \alpha_{15} Z_2 D_3 + \alpha_{16} Z_3 D_2 + \alpha_{17} Z_3 D_3 + \lambda$$

TRT =	Treatment, VR = 1, and zero for SR
Z_i =	1, i = 1, 2, 3, and zero otherwise. Zone 4 is a base zone.
Z_1 =	Zone 1 (Low-potential zone)
Z_2 =	Zone 2 (Medium-potential zone)
Z_3=	Zone 3 (High-potential zone)
Z_4 =	Zone 4 (Very high-potential zone)
Di=	1, i = 2, 3, and Year 1 is the base year
D_1 =	Year 1
D_2 =	Year 2
D_3 =	Year 3
λ =	Lambda (Spatial error coefficient)

Overall base variables: D_1 & Z_4

5.2.1 Diagnostic tests for the Baseline Model

Diagnostic tests on the OLS residuals determine the presence of spatial effects and verify the optimal model. The specification tests on spatial autocorrelation and heteroscedasticity (structural change) are acquired by running the OLS model in conjunction with the weight matrix, which also suggests which model should be used – either the spatial error or spatial lag. The specification tests are then followed by estimating models that incorporate spatial autocorrelation and heteroscedasticity.

Some of the assumptions of the classical linear regression models are that random variables are normally distributed and homoscedastic, and that there is no autocorrelation and multicollinearity between the variables. It was essential to determine whether these conditions hold before estimating the regression coefficients of data collected in this study. Various diagnostic tests for normality, multicollinearity and heteroscedasticity are summarised in Table 5.2.

TABLE 5.2: DIAGNOSTIC TESTS FOR NORMALITY AND HETEROSCEDASTICITY

Test	Pooled data Baseline Model	Year 1 data Baseline Model	Year 2 data Baseline Model	Year 3 data Baseline Model
Jarque-Bera (JB)	1225*	162*	128*	43*
Breusch-Pegan (BP)	202* 254**	81* 41**	48* 43**	33* 32**

*OLS model (JB computed in the OLS model only) **SER model

Normality in the error terms is determined by the Jarque-Bera (JB) test, which evaluates the hypothesis that the residuals are normally distributed. For a normal distribution, skewness should be close to zero and kurtosis closer to three. In the test for normality of errors in the Baseline Model using OLS, a JB value of 1 225 is obtained, based on the data for all the three years aggregated together. The JB value is significant at 1% probability level. JB values of 162, 128 and 43[1] are obtained for the first, second and third years' data,

[1] JB decreased substantially in the third year due to improved management in data collection.

and are also significant at 1% probability level. This means that, if the null hypothesis of normally distributed errors was true, the probability of obtaining such high JB values would be one, indicating non-normality of the error terms.

Heteroscedasticity is mostly common and expected in cross-sectional data, such as the data collected for this study. The Breusch-Pagan (BP) test is a diagnostic test of a regression to determine the presence of heteroscedasticity in the error terms. It is a Lagrange Multiplier test of the hypothesis that the independent variables have no explanatory power on the residual squares (ei2's). The larger the BP test, the greater the evidence against homoscedasticity. The high BP test values calculated in the OLS and the SER models reject the null hypothesis of homoscedasticity, based on the 1% level of significance. A BP value of 202 is obtained in the OLS model for the aggregated data, while the SER model produced a BP value of 254. The BP values under the OLS model for each of the three years are 81, 48 and 33, whereas values of 41, 43 and 32 are correspondingly obtained in the SER models. If the null hypothesis of homoscedasticity was true, the probability of obtaining such high BP values would be one.

The Koenker-Bassett (KB) test in the OLS model also confirms the presence of heteroscedasticity for the aggregated data, as well as for each individual year. A KB value of 82, which is significant at 1% level, was computed for the three years' data aggregated together. The OLS parameter estimates are still unbiased in the presence of heteroscedasticity, but the standard errors are no longer efficient. As a result, inferences from the standard errors may be misleading (Gujarati, 2003). Heteroscedasticity is still persistent in the SER model. Dealing with heteroscedasticity in spatial data presents a problem, as no standard procedure has yet been developed in this regard. Multicollinearity of the models is also tested by GeoDa™. With the exception of the pooled data, all the multicollinearity condition numbers are lower than 20, the recommended maximum condition number

Spatial autocorrelation was also expected in this type of data, and its presence was identified in exploratory data analysis using the univariate Moran's I statistics. As a result, both the OLS and the SER models were estimated, the latter to account for detected spatial autocorrelation. The likelihood ratio test is calculated as a diagnostic for spatial dependence in a SER model, and is significant at 1%.

Due to the presence of spatial variability in the data set and the diagnosis of spatial autocorrelation and spatial heteroscedasticity, the spatial model was more appropriate for analysis so that the spatial effects could be taken into account. However, an OLS regression test had to be conducted for the purpose of spatial diagnostics on the OLS residuals, in order to determine which spatial regression model (lag or error) would be appropriate.

Five diagnostic tests for spatial dependence are reported with the OLS regression output in GeoDa™. These include the Moran's I for the spatial error model, the Lagrange Multiplier (LM) and its robust term for each of the spatial models – the lag and the error models. Based on the spatial diagnostics using the LM values, the spatial error model appears to be the most appropriate for the aggregated, second and third years' data, as demonstrated in Table 5.3.

TABLE 5.3: DIAGNOSTIC TESTS FOR SPATIAL DEPENDENCE

Test	**Value**			
	Aggregate	**Year 1**	**Year 2**	**Year 3**
Moran's I (error)	46*	17*	36*	36*
Lagrange Multiplier (lag)	1 292*	326*	1 206*	883*
Robust LM (lag)	3**	43*	9*	0.08
Lagrange Multiplier (error)	1 580*	284*	1 213*	1 267*
Robust LM (error)	291*	0.48	16*	383*

*****Significant at 1% **Significant at 10%

The model with the highest LM value and its robust term is the most appropriate. The spatial lag model is reflected as the most appropriate for the first year's data. The spatial lag model relates to spatial autocorrelation in the dependent variable (y) in the model itself, which may be lagged in time and/or space. The mechanism for spatial lag models in crop data is hard to fathom. It is difficult to grasp that the yield in one part of the field is influenced by yield in another part of the field, or that yield obtained in the previous season has an effect on yield in the current season. The spatial error model is selected for year one, based on Griffin *et al.* (2005b).

According to Griffin *et al.* (2005b), the spatial error model is usually appropriate for large-scale on-farm data**.** It is appropriate when the spatial structure is captured in the residuals of the regression, or all the omitted variables are included in the residual. Many factors affect yield variability at field level, and it is not possible to include all the variables in the model. The variables that cannot be included in the model are captured in the residual, making the spatial error model the most suitable model (Griffin *et al.*, 2005).

5.2.2 Baseline Model: The Treatment Model

The results of coefficient estimates for the Baseline Model estimated with the OLS and SER models for the three years aggregated together are summarised in Table 5.4, while the regression outputs for the individual years are included as Tables 5.1 to 5.6 in Annexure 5. The results of the data aggregated together (pooled Baseline Model) are discussed in this section, while the results of the year-by-year Baseline Models are discussed in Section 5.2.6.

TABLE 5.4: MEASURES OF FIT AND COEFFICIENT ESTIMATES WITH REGARD TO THE POOLED DATA

Measures of fit	OLS	SER
R-squared	0.59	0.80
Adjusted R-squared	0.58	-
F-statistic/Likelihood Ratio Test	183	846
Prob (*F*-statistic)	0.00	0.00
Log-likelihood	-2 153	-1 729
Akaike Information Criterion	4 341	3 495
Schwartz Criterion	4 443	3 597

Regression output								
	OLS				SER			
Variable	Coeffi-cient	Std Error	t-Statis-tics	Proba-bility	Coeffi-cient	Std Error	z-Statis-tics	Prob-ability
Constant	4.6139	0.0953	48.4113	0.0000	4.5011	0.0921	48.8617	0.0000
TRT	0.3864	0.1028	3.7579	0.0002	0.3193	0.0707	4.5183	0.0000
Z_1	0.3395	0.1136	2.9890	0.0028	0.4007	0.1109	3.6120	0.0003
Z_2	0.1218	0.1016	1.1985	0.2309	0.2899	0.0987	2.9355	0.0033
Z_3	0.1810	0.1111	1.6296	0.1033	0.2686	0.1069	2.5126	0.0120
TRT_Z_1	-0.4130	0.1138	-3.6286	0.0003	-0.3763	0.0782	-4.8109	0.0000
TRT_Z_2	-0.4382	0.1024	-4.2811	0.0000	-0.3633	0.0704	-5.1640	0.0000
TRT_Z_3	-0.3053	0.1120	-2.7263	0.0065	-0.2467	0.0770	-3.2024	0.0014
D_2	0.0794	0.1207	0.6576	0.5108	0.0089	0.0832	0.1064	0.9153
D_3	0.4337	0.1207	3.5937	0.0003	0.3772	0.0833	4.5267	0.0000
D_2_TRT	0.7692	0.0688	11.1738	0.0000	0.7956	0.0470	16.9113	0.0000
D_3_TRT	0.0605	0.0687	0.8809	0.3785	0.0569	0.0469	1.2120	0.2255
Z_1_D_2	-0.0949	0.1402	-0.6770	0.4985	-0.0406	0.0965	-0.4204	0.6742
Z_1_D_3	1.1555	0.1397	8.2705	0.0000	1.2471	0.0965	12.9268	0.0000
Z_2_D_2	-0.0680	0.1260	-0.5397	0.5895	-0.0223	0.0868	-0.2564	0.7976
Z_2_D_3	1.2416	0.1261	9.8503	0.0000	1.3043	0.0872	14.9640	0.0000
Z_3_D_2	0.0482	0.1378	0.3500	0.7264	0.1116	0.0949	1.1764	0.2394
Z_3_D_3	1.5740	0.1377	11.4323	0.0000	1.6468	0.0950	17.3269	0.0000
LAMBDA					0.6260	0.0121	51.8332	0.0000

5.2.2.1 Model fit

The R^2 is usually used as a measure of goodness of fit in ordinary regression analysis, with the model having the highest R^2 considered to have the best fit. The best fit implies that the predicted values match the observed values for the dependent variable. However, over-fitting occurs with the R^2, as it increases in value with additional explanatory variables. The adjusted R^2 corrects for this over-fitting since it does not increase as the number of explanatory variables increases, but the adjusted R^2 is not reported in GeoDa™.

In Table 5.4, the R^2 obtained in the SER model is higher at 80% than the 59% obtained with the OLS model. The R^2 of 80% for the aggregated data means that the explanatory variables explain 80% of the variation in yield. Model fit statistics for each of the three years are presented in Table 5.5.

TABLE 5.5: MEASURES OF GOODNESS OF FIT FOR EACH OF THE THREE YEARS

Measures of goodness of fit	Year 1	Year 2	Year 3
R-Squared (R^2)			
OLS	2%	27%	36%
SER	78%	63%	82%
Log-likelihood			
OLS	-714	-738	-624
SER	-400	-532	-225
Akaike Information Criterion (AIC)			
OLS	1 441	1 492	1 264
SER	817	1 080	466
Schwartz Criterion (SC)			
OLS	1 450	1 353	1 007
SER	1 480	1 529	1 300

The R^2 values of 78%, 63% and 82% are obtained for the first, second and third years in that order for the SER models, relative to the R^2 values of 2% for the first year, 27% for the second year and 36% for the third year obtained with the OLS models. Looking at the R^2, it appears that the SER models display superior performance in comparison with the OLS models. However, Gujarati (2003) argues that it is not appropriate to judge the goodness of fit of the model based on R^2 alone, as it does not indicate whether the estimated partial regression coefficients are statistically different from zero. Some of them may be different, while others may not be.

Nevertheless, R^2 is not appropriate as a measure of fit in comparing spatial regression models, and the Maximum Likelihood-based models become more reliable. GeoDa™ reports the logarithm of the likelihood obtained for the OLS estimates. This likelihood can serve as an alternative to the R^2 in measuring the goodness of fit of a model. The log-likelihood value measures how good or poorly the model predicts the output in the observed data. The model with the highest log-likelihood has the best fit.

The SER regression models have higher log-likelihood values in comparison to the OLS models in the pooled data, as well as in all the individual years (Table 5.5). In the pooled data, a log-likelihood value of -2 153 is obtained with the OLS, in comparison with the log-likelihood value of -1 730 obtained with the SER model. Log-likelihood values of -714, -738 and -624, respectively, are obtained for each of the three years with the OLS, in comparison to higher log-likelihood values of -400 in the first year, -532 in the second year and -225 in the third year with the SER model. This means that estimated parameters make the observed data most likely with the SER model, since the higher values of the log-likelihood statistic indicate the more desirable model (SAS Institute Inc., 1999). However, the log-likelihood increases with additional variables, as does the R^2, over-fitting the model. This over-fitting can be corrected by employing the Information Criteria, the Akaike Information Criterion (AIC) and the Schwartz Criterion (SC). The lower the Information Criteria value, the better the model.

In Table 5.5, the AIC constitutes the other decisive factor in selecting the optimal model. Unlike in the case of the log-likelihood criterion, where a preference is exercised for bigger statistics, the smaller the AIC value is, the better the fit of the model to the observed data will be. The AIC value of 3 495 calculated in the SER model for the pooled data is smaller than the AIC value of 4 341 computed with the OLS model. The AIC values of 1 444, 1 492 and 1 264, respectively, for the first, second and third years are calculated in the OLS model, relative to smaller AIC values of 817, 1 080, and 466 in the SER model. The SER model therefore provides the closest fit approximating the true model. The SC works in the same way as the AIC regarding preference for the smaller Schwartz value. The SER model is still more appropriate than the OLS model when considering their Schwartz values, which are 1 480, 1 529 and 1 300, respectively, for the SER models, in comparison with 1 450, 1 353 and 1 007 for the OLS models in each year. In the pooled data, the Schwartz value of 3 597 achieved with the SER model is smaller than that obtained with the OLS model (4 443).

Another measure of fit in spatial regression models is the Likelihood Ratio Test. In the OLS Regression, the likelihood ratio statistic is analogous to an overall *F*-test. The *F*-statistics calculated under the OLS model of 183 for the pooled data, including 2.55, 38 and 58 for each of the years in chronological order, are highly significant at 1%, indicating that the explanatory variables, effective depth (ED), N, Zone 1 (Z_1), Zone 2 (Z_2), and Zone 3 (Z_3), as well as treatment (TRT), simultaneously have an effect on yield. The likelihood ratio tests computed under the SER model are bigger than the *F*-test, and significant for all the data sets. Nevertheless, caution should be exercised in using the *F*-test/likelihood ratio test as a criterion for model selection, since – in comparison with the AIC and the SC – it has a tendency to select simpler models, even in cases where the complex model may be more appropriate (Ludden, Beal & Sheiner, 1994).

5.2.2.2 Regression coefficient estimates: The aggregated data

Spatial econometric models require specification of the weights matrix. In this study, the definition of neighbours is based on the geographic criteria in terms of the distance between yield points, using the Cartesian space (longitude and latitude). The Euclidean distance-based matrices were calculated in GeoDa™ using the threshold (minimum) distances for each year, which ensures that each observation has at least one neighbour. (Each observation had more than one neighbour in this case.) The maps were first projected into metres in ArcView GIS 3.2 to allow usage of distance-based matrices. A SER regression model that considers autocorrelation of the residuals was subsequently estimated, using a Baseline Model according to different management zones. The year dummy variables were created when the data of the three study years were aggregated together in a single regression. The first year (2001/2002) was used as a base year. The regression results of pooled data for the three years are discussed first, culminating in the results of the models of individual years. Although the regression outputs for OLS models are included for comparison purposes, only the results obtained with SER models are discussed, in view of the limitations of the OLS model discussed earlier. The coefficient estimates of the OLS model are unreliable, and may be misleading.

The significance test on the coefficients of dummy variables, Z_1, Z_2, and Z_3, assess the yield variation according to management zone. Zone 4 (Z_4), which is the very high-potential zone, serves as the base, with other management zones deviating from Z_4.

With the Baseline Model in Table 5.4, the zone dummy variable is defined as $Z_i = 1$, $i = 1, 2, 3$ and zero otherwise, with Zone 4 used as the base. There is only a slight variation in the coefficient estimates of the SER model in comparison with the OLS model. With the spatial model, all the estimates for the zones are statistically significant at 1% level, indicating that there are coefficient differences between the zones. As the spatial error coefficient (Lambda) is highly significant, the SER model is expected to improve the model fit by taking the spatial effects into account.

The year dummy variable for the second year and its interaction with all the zones, as well as year three and the treatment interactions, are not statistically significant. The rest of the coefficients are significant at 1% level, connoting that yields in these zones are significantly different from the yields in the very high-potential zone (Z_4), the base zone. The coefficients of all the zones have positive signs, while their interactions with the treatment have negative signs.

The positive coefficients for the zones imply that average yields in these zones are higher than average yields obtained in the base zone. With reference to the descriptive statistics, this applies to Zones 1 and 3 in year two, and to all the zones in year three. The positive TRT coefficient indicates that the VR treatment has a positive effect, mainly on Zone 4. The TRT coefficient of 0.32 suggests that the VR treatment in Zone 4 results in yields 0.32 kg higher than the yields produced by the SR treatment in the same zone. This is consistent with descriptive statistics, as the VR treatment in Zone 4 has been performing better than the SR treatment over all the years. The effect of the VR application treatment on each zone in each year, as reflected by the slopes, is summarised in Table 5.6.

TABLE 5.6: EFFECT OF VARIABLE-RATE TREATMENT ON EXPECTED MAIZE YIELDS

	Zone 1		Zone 2		Zone 3		Zone 4	
	SR Yield	VR Yield Advan-tage	SR Yield	VR Yield Adva-ntage	SR Yield	VR Yield Advan-tage	SR Yield	VR Yield Advan-tage
Year 1	4.90	-0.06	4.79	-0.04	4.77	0.07	4.50	0.32
Year 2	4.87	0.74	4.78	0.75	4.89	0.87	4.51	1.11
Year 3	6.53	-0.0001	6.47	0.01	6.79	0.013	4.88	0.38

The VR treatment effect in Zones 1, 2 and 3 is close to zero in years one and three, while in year two the treatment effect is positive in all the zones. The VR effect is substantial and positive in Zone 4 in all the years. The negative VR effect in Zone 1 in years one and three, and in Zone 2 in year three, is very small and statistically not significantly different from the SR yields.

The Baseline Model is used in Section 5.3 to calculate the profitability of variable-rate technology (VRT) as a package, in comparison with the uniform application (SR). The cost of a wrong decision, which is captured by using the OLS (wrong model) instead of the SER model, is determined by comparing the profit based on the two models. As this is a linear additive model, the assumption is that the effect of the treatment (TRT) on yield takes place through intercept changes.

Before the results of the sensitivity tests are presented, a comparison of the model fit between the SER models is summarised in Table 5.7.

TABLE 5.7: MODEL FIT BETWEEN THE MODELS

	Baseline Model	Sensitivity Test 1	Sensitivity Test 2	Sensitivity Test 3
Adjusted R^2	0.80	0.76	0.79	0.83
F_Test/ Likelihood Ratio Test	846	554	692	677
Log- likelihood	-1 729	-1 844	-1 758	-1 542
AIC	3 495	3 715	3 565	3 146
SC	3 597	3 789	3 702	3 322

Because it models both the zones and the continuous N rate, Sensitivity Test 3 appears to have a slightly better fit than the other models, but it also suffers from extreme multicollinearity. One variable (TRT_N^2) was omitted from the Sensitivity Test 3 model, as it caused almost perfect multicollinearity and prevented estimation. The Baseline Model, which had the second-best fit, was selected for further analysis since it was the model assumed in the original design of the on-farm trial, and provides a robust analysis with much less multicollinearity.

5.2.3 Sensitivity Test 1: The TRT Effective-depth Model

Sensitivity Test 1, the model that comprises the treatment and effective depth as explanatory variables, seeks to determine whether the treatment varies with the effective depth, and thus differs between the zones. The results of this model are consistent with the results of the Baseline Model. Table 5.8 presents the results of Sensitivity Test 1, as estimated by both the OLS and the SER models.

TABLE 5.8: COEFFICIENT ESTIMATES FOR SENSITIVITY TEST 1 (ED_TRT MODEL)

	OLS				SER			
Variable	Coeffi-cient	Std Error	z-value	Proba-bility	Coeffi-cient	Std Error	z-value	Proba-bility
Constant	1.4257	0.2981	4.7825	0.0000	1.47853	0.2935	5.0378	0.0000
TRT	-0.2499	0.1309	-1.9095	0.0563	-0.28085	0.0999	-2.8122	0.0049
ED_(cm)	0.0361	0.0036	9.9743	0.0000	0.04	0.0036	10.1280	0.0000
$ED\_^2$	-0.0001	0.0000	-8.3938	0.0000	-0.00009	0.0000	-8.6470	0.0000
ED_TRT	0.0014	0.0007	2.0647	0.0391	0.002	0.0005	3.0280	0.0025
D_2	0.0806	0.4901	0.1645	0.8694	-0.37778	0.3951	-0.9562	0.3389
D_3	4.0343	0.3834	10.5233	0.0000	3.66962	0.3071	11.9498	0.0000
D_2_TRT	0.7888	0.0678	11.6422	0.0000	0.80100	0.0515	15.5514	0.0000
D_3_TRT	0.0790	0.0676	1.1691	0.2425	0.07529	0.0514	1.4646	0.1430
D_2_ED	-0.0019	0.0059	-0.3274	0.7434	0.00315	0.0048	0.6584	0.5103
D_3_ED	-0.0290	0.0047	-6.2008	0.0000	-0.02523	0.0037	-6.7408	0.0000
$D_2\_ED^2$	0.0000	0.0000	0.4937	0.6216	-0.00001	0.0000	-0.3722	0.7098
$D_3\_ED^2$	0.0001	0.0000	5.9067	0.0000	0.00007	0.0000	6.6024	0.0000
LAMBDA					0.54101	0.0148	36.5258	0.0000

Measures of fit	OLS	SER
Adjusted R^2	0.60	0.76
F_Test/ Likelihood Ratio Test	272 (0.00)	554 (0.00)
Log-likelihood	-2121	-1844
Akaike Information Criterion	4269	3715
Schwarts Criterion	4342	3789

There is an improvement in the model fit from the OLS to the SER model, as is evident from the increase in the log-likelihood value, from -2 121 in the OLS model to -1 844 in the SER model. The information criteria, AIC and SC, also indicate this improvement in the model fit. There is a decline in the AIC value from 4 269 in OLS to 3 715 in the SER model, while a decrease from 4 342 in the OLS to 3 789 is observed in the SC.

The coefficients for the effective depth (ED and ED^2) have the expected signs, and are significant. The effective depth conforms to the expectation of a positive effect on yield, as it has a positive sign. The ED coefficient of 0.036 means that, on average, there is an increase in yield of about 0.04 tons/ha for each centimetre (cm) increase in depth to a certain extent on this field. The depth of the soil is related to its ability to store moisture, and the availability thereof to the plants. The combined effects of the TRT and ED_TRT variables indicate that the treatment effect in year one was positive for soils greater than 173 cm. As the average soil depth in the field is 174 cm, this means that, in year one, VR had a positive effect in this field for soils over the average depth.

The Sensitivity Test 1 model is used to estimate yields for both treatments. Table 5.9 shows the expected yields with VR and SR treatments, taking into account the effective depth.

TABLE 5.9: EXPECTED YIELDS[1] WITH VR AND SR TREATMENTS USING SENSITIVITY TEST 1 MODEL (tons/ha)

	Zone 1		Zone 2		Zone 3		Zone 4	
	Expected SR yield	VR Advan-tage	Expected SR yield	VR Advan-tage	Expected SR yield	VR Advan-tage	Expected SR yield	VR Advan-tage
Year 1	4.93	-0.01	4.88	-0.02	4.95	0.03	4.93	0.01
Year 2	4.94	0.80	4.87	0.79	4.99	0.83	4.95	0.81
Year 3	6.45	0.07	6.40	0.06	6.53	0.11	6.46	0.08

[1] Yields estimated at the average soil depth for the zone.

The VR treatment has a yield advantage at average zone soil depth for all the zones in years two and three. In year two, that yield advantage is about 0.8 tons/ha in all zones. It can therefore be concluded that treatment does vary by depth and that the effective soil depth can be used as one of the determinants of management zones, but not as the main criterion. Other variables such as the spatial yield distribution of yield across the field over time should also be taken into account.

5.2.4 Sensitivity Test 2: The Nitrogen-Zone Model

To determine whether N response varies by zone, Sensitivity Test 2, the model that includes N and the different zones as the explanatory variables, was estimated. The coefficient estimates for both the OLS and the SER models are presented in Table 5.10.

TABLE 5.10: COEFFICIENT ESTIMATES FOR SENSITIVITY TEST 2 (N-ZONES MODEL)

OLS					**SER**			
Variable	Coefficient	Std Error	t-value	Probability	Coefficient	Std Error	z-value	Probability
Constant	-0.4732	0.4517	-1.0474	0.2950	-0.2218	0.3874	-0.5727	0.5669
N (kg/ha)	-0.0099	0.0169	-0.5843	0.5591	-0.0109	0.0135	-0.8067	0.4199
N^2	0.0006	0.0001	4.1379	0.0000	0.0006	0.0001	5.1736	0.0000
Z_1	12.0131	2.3889	5.0287	0.0000	11.2593	1.8285	6.1577	0.0000
Z_2	6.9536	1.4136	4.9192	0.0000	8.7513	1.0859	8.0592	0.0000
Z_3	-35.945	7.6920	-4.6731	0.0000	-22.616	6.0781	-3.7210	0.0002
N_Z_1	-0.1331	0.0462	-2.8832	0.0040	-0.1236	0.0353	-3.5003	0.0005
N_Z_2	-0.0394	0.0180	-2.1844	0.0290	-0.0724	0.0140	-5.1580	0.0000
N_Z_3	0.8302	0.1546	5.3699	0.0000	0.5676	0.1222	4.6426	0.0000
$N2\_Z_1$	0.0002	0.0002	0.7090	0.4784	0.0001	0.0002	0.8097	0.4181
$N2\_Z_2$	-0.0003	0.0001	-4.0835	0.0000	-0.0001	0.0001	-2.2808	0.0226
$N2\_Z_3$	-0.0047	0.0008	-5.9611	0.0000	-0.0034	0.0006	-5.4479	0.0000
D_2	-10.860	2.3199	-4.6816	0.0000	-9.2179	1.7846	-5.1654	0.0000
D_3	-4.5582	1.4546	-3.1337	0.0017	-4.5846	1.1313	-4.0527	0.0001
D_2_N	0.2917	0.0511	5.7046	0.0000	0.2504	0.0393	6.3692	0.0000
D_3_N	0.0936	0.0298	3.1444	0.0017	0.1002	0.0232	4.3228	0.0000
$D_2\_N^2$	-0.0018	0.0003	-6.3513	0.0000	-0.0015	0.0002	-7.1286	0.0000
$D_3\_N^2$	-0.0005	0.0002	-2.9502	0.0032	-0.0005	0.0001	-4.2828	0.0000
$Z1\_D_2$	-0.4123	0.1512	-2.7265	0.0065	-0.3733	0.1092	-3.4184	0.0006
$Z_1\_D_3$	1.4935	0.1505	9.9233	0.0000	1.4598	0.1090	13.3911	0.0000
$Z_2\_D_2$	-0.3415	0.1351	-2.5280	0.0115	-0.3710	0.0977	-3.7986	0.0001
$Z_2\_D_3$	1.4038	0.1290	10.8798	0.0000	1.4877	0.0936	15.9022	0.0000
$Z_3\_D_2$	-0.1838	0.1429	-1.2862	0.1985	-0.1050	0.1037	-1.0121	0.3115
$Z_3\_D_3$	1.8417	0.1393	13.2235	0.0000	1.8026	0.1010	17.8545	0.0000
LAMBDA	-	-	-	-	0.5939	0.0131	45.3379	0.0000

Measures of fit	**OLS**	**SER**
Adjusted R^2	0.60	0.79
F-Test/ Likelihood Ratio Test	145 (0.00)	-1 758 (0.00)
Log-likelihood	-2 105	692
Akaike Information Criterion	4 258	3 565
Schwarts Criterion	4 394	3 702

* Probabilities in parentheses

The SER model consistently results in an improved fit, as indicated by the measures of fit. In Sensitivity Test 2, the N coefficient is negative but insignificant, while its quadratic form is positive and significant in year one. The estimated model is therefore convex for N, with a minimum very close to N=zero, since N*=9.45 kg/ha in year one. In years two and three, the estimated N model is concave, as expected, since the years by N dummies (D_N) are positive while the quadratic N by year dummies (D_N^2) are negative. As expected, the N-zone cross terms are significant, as the N application was varied deliberately between the zones. The interactions between the zones and the years are statistically significant, indicating that yields do vary over the years and thus confirming the results obtained in Chapter 4. The yield-maximising N rates for each zone and year, computed on the basis of the coefficient estimates of Sensitivity Test 2, are summarised in Table 5.11.

TABLE 5.11: YIELD-MAXIMISING NITROGEN RATES (N)

Crop Season	Zone 1	Zone 2	Zone 3	Zone 4
2002/2003				
N*[1]	92.95	91.80	99.70	9.45
SR N	105*	105*	105*	105*
VR N	84**	92**	100**	110**
2003/2004				
N*[1]	72.19	77.91	93.46	125.91
SR N	102*	102*	102*	102*
VR N	69**	85**	110**	123**
2004/2005				
N*[1]	84.54	126.21	99.15	40.77
SR N	102*	102*	102*	102*
VR N	69**	85**	114**	127**

[1]N*= Yield-maximising N rate from the Nitrogen-Zone Model (Sensitivity Test 2). SR N* = rate applied in the SR treatment, and VR N** = rate applied in the zone for the VR treatment.

There is a huge variation in yield-maximising N rates for Zone 4 in the three study years, while the yield-maximising N levels for Zone 3 are virtually identical. In comparing the yield-maximising N rates with the N applications used at the trials, there is not much difference for VR treatment, except in Zone 4. The substantial difference between the yield-maximising N rates and the trial rates occurs in Zone 4. Apart from the application for Zone 2 in year three and Zone 4 in year two, the constant applications (SR treatment rates) are higher than the yield-maximising rates. The highest yield-maximising N level is found in Zone 2 for the third year, while the lowest occurs in year one for Zone 4. Zone 2 happens to be the lowest-producing zone in the first two years, while Zone 4 was the highest in year one.

Because N is applied variably, the varying yield-maximising N rates in different zones indicate that the crop response to N will vary by zone. This is consistent with the Baseline Model, which indicated that the treatment effect varies across the zones. The finding that the VR N rates used at the trials approximate the profit-maximising N rates, except in Zone 4, suggests that economic analysis of the trial results should give a good indication of the profit potential of VR. Profit analysis is discussed in Section 5.3.

5.2.5 Sensitivity Test 3: The Nitrogen-Zone-Treatment Model

Even though the analysis based on Sensitivity Test 2, the Nitrogen-Zone Model, evaluates the performance within the zones and between the years, it does not distinguish between the two treatments. The same model is explored further, except that the TRT variable and its interaction with the zones, the years and the linear N are included. The interaction between the TRT and the quadratic N is excluded, as it causes extreme multicollinearity. The coefficient estimates of this model are presented in Table 5.12.

TABLE 5.12: COEFFICIENT ESTIMATES FOR SENSITIVITY TEST 2 (NITROGEN-ZONE-TRT MODEL)

Variable	Coefficient	Std. Error	z-value	Probability
Constant	0.1340	0.3514	0.3815	0.7029
TRT	-42.9125	4.7279	-9.0765	0.0000
N_(kg/ha)	0.5193	0.0563	9.2209	0.0000
$N\_^2$	-0.0045	0.0005	-8.4501	0.0000
Z_1	4.1622	1.9007	2.1898	0.0285
Z_2	4.6567	1.3335	3.4921	0.0005
Z_3	-15.4235	5.6878	-2.7117	0.0067
N_Z_1	-0.0333	0.0371	-0.8961	0.3702
N_Z_2	-0.0375	0.0159	-2.3617	0.0182
N_Z_3	0.3178	0.1132	2.8076	0.0050
$N^2\_Z_1$	-0.0001	0.0002	-0.2807	0.7789
$N^2\_Z_2$	-0.0001	0.0000	-1.7406	0.0818
$N^2\_Z_3$	-0.0016	0.0006	-2.8830	0.0039
TRT_Z_1	0.7384	0.2650	2.7868	0.0053
TRT_Z_2	0.0347	0.1025	0.3386	0.7349
TRT_Z_3	0.3118	0.1049	2.9716	0.0030
N_TRT	0.4085	0.0446	9.1490	0.0000
D_2	36.6736	4.5398	8.0783	0.0000
D_3	40.3859	4.3771	9.2267	0.0000
D_2_N	-0.7984	0.1021	-7.8201	0.0000
D_3_N	-0.8671	0.0989	-8.7690	0.0000
$D_2\_N^2$	0.0041	0.0005	7.5438	0.0000
$D_3\_N^2$	0.0045	0.0005	8.4426	0.0000
D_2_TRT	2.1349	0.1983	10.7675	0.0000
D_3_TRT	1.2462	0.2042	6.1036	0.0000
$Z_1\_D_2$	0.1702	0.1218	1.3979	0.1621
$Z_1\_D_3$	1.3123	0.1174	11.1818	0.0000
$Z_2\_D_2$	0.2610	0.1020	2.5597	0.0105
$Z_2\_D_3$	1.4709	0.0969	15.1735	0.0000
$Z_3\_D_2$	0.3225	0.0996	3.2374	0.0012
$Z_3\_D_3$	1.9980	0.0995	20.0813	0.0000
LAMBDA	0.6047	0.0128	47.3981	0.0000

Measures of fit	Sensitivity Test 2	Sensitivity Test 3
Adjusted R^2	0.79	0.82
F_Test/ Likelihood Ratio Test	-1 758 (0.00)	-1 542
Log-likelihood	692	676
Akaike Information Criterion (AIC)	3 565	3 146
Schwarts Criterion (SC)	3 702	3 322

The inclusion of the TRT variable improves the model fit, as indicated by the increase in the log-likelihood from -1 758 to -1 542 from Sensitivity Test 2 to Sensitivity Test 3, and a decrease in AIC and the SC from 3 565 and 3 702 to 3 146 and 3 322, respectively.

In contrast to Sensitivity Test 2, in which only year one showed a convex response to N, the N response in Sensitivity Test 3 is convex for all three years. Apart from Zone 1, all the interactions between the N and the zones are significant, supporting the earlier conclusion that N response does vary between the zones. The results of this model further reinforce the results of the Baseline Model, which indicated that the treatment effect does vary between zones and years.

As TRT is essential in calculating profitability and is regarded as one of the most important variables, it was essential to determine whether TRT is significant over the entire field. This was determined by the Likelihood Ratio Test. The likelihood ratio statistic follows a chi-square distribution, with the degrees of freedom equalling the difference in the number of variables between the two models. Under the Maximum Likelihood estimation, the null hypothesis is that the addition of TRT and its interaction terms to Sensitivity Test 2 does not result in a significantly improved fit – i.e. Ho: $\Theta = 0$. If this hypothesis is true and the restriction $\Theta = 0$ is valid, omission of these variables should not lead to a large reduction in the log-likelihood function.

Sensitivity Test 3 is regarded as an unrestricted model (additional parameters), whilst Sensitivity Test 2 is restricted. The log-likelihood value of the unrestricted model is L_U (the value of the likelihood function at the unconstrained value of Θ), and L_R represents the likelihood value of the restricted model. If the likelihood ratio statistic λ exceeds the appropriate critical value from the chi-square table, the null hypothesis is rejected. Sensitivity Test 3 has 10 more parameters than Sensitivity Test 2, and this was used as the degrees of freedom. With 10 degrees of freedom and a 5% significance level, the calculated likelihood ratio value of 936 is greater than the value of the test statistic (25), and it is therefore concluded that TRT is significantly different from zero. It was therefore concluded that TRT is important as a variable to calculate the profitability of VRT.

5.2.6 Regression coefficient estimates: Single-year Baseline Models

The analysis of the Single-year Baseline Models aims to identify the consistency with or deviations from the Baseline results of the aggregated data. In year one, with the exception of all the three zones, all the estimated partial regression coefficients in the Baseline Model are individually statistically significant, considering their very low probabilities. The TRT is positive, while all the TRT-zone interactions are negative. The positive TRT implies that, in year one, VR treatment performed better than the SR treatment. However, according to the descriptive statistics, the yields obtained under the two treatments are the same. This shows that inaccurate decisions can be made by using only normal arithmetic means and not conducting regression analysis, which takes spatial associations between variables into account. In the Baseline Model for the pooled data, the TRT coefficient is positive and significant, consistent with each individual year. Table 5.13 presents a summary of the regression coefficient estimates for the Baseline Model of each individual year.

TABLE 5.13: BASELINE REGRESSION ESTIMATES FOR EACH YEAR

Variable	Coefficient		
	Year 1	Year 2	Year 3
Constant	4.74* (65.37)	4.84* (31.48)	4.6* (32.19)
TRT	0.13* (2.54)	0.79* (6.95)	0.78* (11.80)
Z_1	-0.02 (-0.19)	-0.01 (-0.12)	1.81* (23.98)
Z_2	-0.02 (-0.22)	-0.05 (-0.53)	1.83* (27.76)
Z_3	0.04 (0.56)	-0.09 (-0.78)	1.90* (26.01)
TRT_Z_1	-0.11*** (-1.72)	-0.07 (-0.49)	-0.78* (-9.65)
TRT_Z_2	-0.18* (-3.18)	-0.06 (-0.47)	-0.79* (-10.91)
TRT_Z_3	-0.14** (-2.17)	0.27** (1.99)	-0.79* (-9.94)
LAMBDA	0.66 (44.20)	0.86 (28.02)	0.92* (55.04)

Z-values in parentheses *1%, **5% and ***10% levels of significance

According to Table 5.13, the coefficients of all the zones are insignificant in year one, but their interactions with TRT are significant, implying that the treatment varies by zone. All the zones and their interaction with TRT, excluding the interaction between TRT and Zone 3, are insignificant in the year two Baseline Model. TRT remains positive and significant, as in the pooled Baseline Model. The positive TRT in year two coincides with descriptive statistics, as VR resulted in higher average yields in all the zones in year two. From the insignificant interaction terms it can be concluded that, even though VR out-performed SR in year two, yields under the VR treatment did not vary much between the zones, but VR yield advantage occurred in all the zones. In support of this, the null hypotheses of no significant differences between the mean yields of the four zones were all rejected in the t-tests of descriptive analysis.

A different picture emerges in year three. All the zones and their interaction terms with TRT are significant. The coefficients for all the zones are positive, substantiating the descriptive statistics, as Zone 4 had the lowest performance in year three. TRT is positive and significant, meaning that there is a significant difference between yields under VR treatment and yields under SR treatment. This is still in line with the results obtained with the pooled Sensitivity Test 2 Model. However, the yields from the two treatments vary modestly, except for Zone 4, where yields for the VR exceeded those obtained for SR by 780 kg/ha.

In each of the models for Sensitivity Test 1, all the TRT coefficients are negative and insignificant, but the ED_TRT interactions are positive. The conclusion generated from the aggregated data in Sensitivity Test 1, which stated that the treatment effect is positive for soils over the average depth (174 cm) in year one, and in shallower soils in years two and three, is sustained in each individual year model. The effect of the VR treatment on expected yields, as estimated on the basis of individual year models, is summarised in Table 5.14.

TABLE 5.14: THE EFFECT OF THE VR TREATMENT ON MAIZE YIELDS – INDIVIDUAL YEAR ANALYSIS

	Zone 1		Zone 2		Zone 3		Zone 4	
	SR Yield	VR Yield Advan-tage	SR Yield	VR Yield Advan-tage	SR Yield	VR Yield Advan-tage	SR Yield	VR Yield Advan-tage
Year 1	4.72	0.026	4.72	-0.047	4.78	-0.047	4.74	0.135
Year 2	4.83	0.718	4.79	0.733	4.76	0.699	4.84	0.786
Year 3	6.47	0.009	6.49	-0.005	6.56	0.000	4.66	0.788

The effect of the VR treatment on individual years is consistent with the results of the aggregated data in the Baseline Model analysis (Table 5.6). In year two, the treatment effect is large and positive in all the zones; the VR effect is considerable and positive in Zone 4 in all the years. The negative VR effect in Zone 2 in years one and three, and in Zone 3 in year one, is very small and statistically not significantly different from the SR yields. In terms of VR yield advantage, the individual years and aggregate analyses differ from the effective soil depth model results only for Zone 4 (Table 5.9). The effective soil depth model shows positive, but relatively small benefits for VR in Zone 4 in years one and three. Of the four zones, Zone 4 was expected to show the highest yields. However, this has not been the case. The improper identification of management zones, which resulted in Zone 4 being classified as the very high-potential zone, while, in fact, it includes average and some low areas, caused this unexpected behaviour of the zone.

5.2.7 Regression coefficient estimates: Individual years Sensitivity Test 2 Models

As with the pooled data model, the Sensitivity Test 2 for each year shows that N application varies according to zone. The entire cross terms between N and the zones are significant at 5%, implying that there is a significant difference in N application in these zones relative to the overall mean response. The N responses for each year are consistent with the pooled Sensitivity Test 2. The first year exhibits the convex production function, while the second and third years depict a normal concave agricultural production function with a positive linear coefficient and a negative quadratic coefficient, thus upholding the law of diminishing marginal returns. However, both the linear and quadratic forms of N are not significant in year two, while its quadratic form is not significant in the third year.

The distribution of the N coefficient estimates was simulated to portray the maize crop marginal response to N according to zone, using the SER model regression output in Table 5.2. The mean N effect and the Nitrogen-Zone interaction terms were used in the computation of the marginal yield. The expected yield was calculated by varying the N-Zone coefficients and adjusting the N rate by an increment of 5 kg/ha. In assessing the response per zone, only Zone 3 exhibits the normal concave production in each of the three years, as indicated by Figures 5.1 to 5.3.

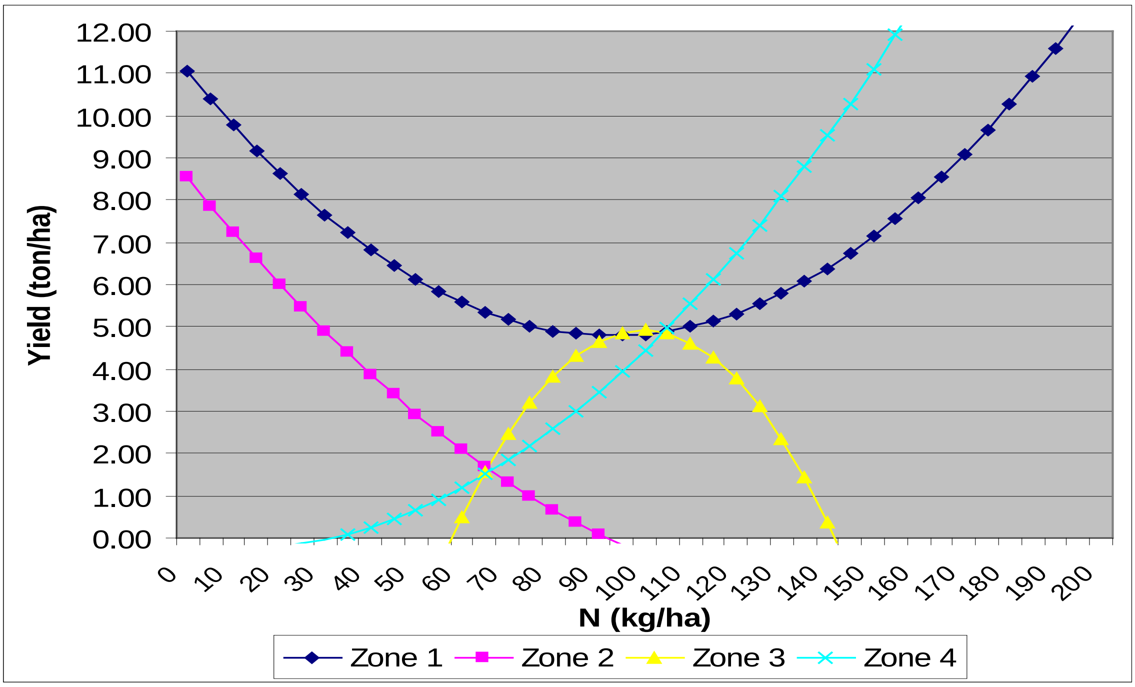

Figure 5.1: Crop response to N according to zone – Year 1 (2002/2003)

Zones 1, 2 and 4 display convex production curves (Figure 5.1); hence, an average of the four functions results in a convex production for the overall year.

The maize crop marginal response to N according to zone for year two (2002/2003) is presented in Figure 5.2.

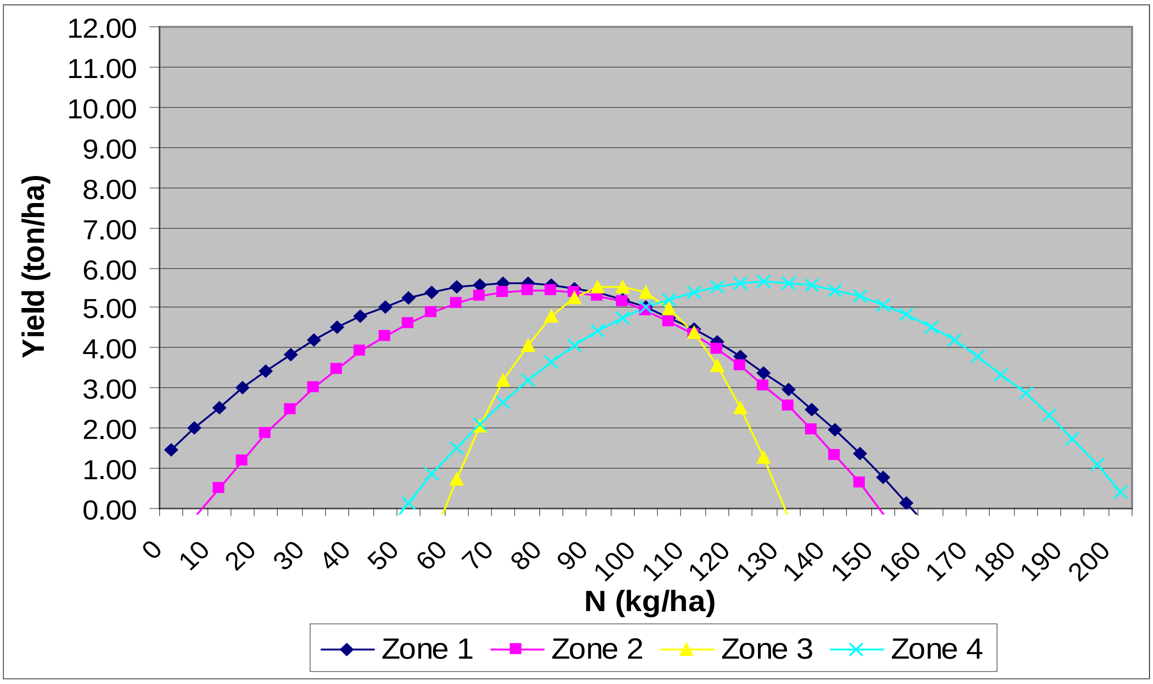

Figure 5.4: Crop response to N according to zone – Year 2 (2003/2004)

As illustrated in Figure 5.2, the production functions for all the zones are concave in year two. In year three, Zone 3 still predominately displays a concave curve, while the rest do not.

Figure 5.3 illustrates the N response for year three (2004/2005).

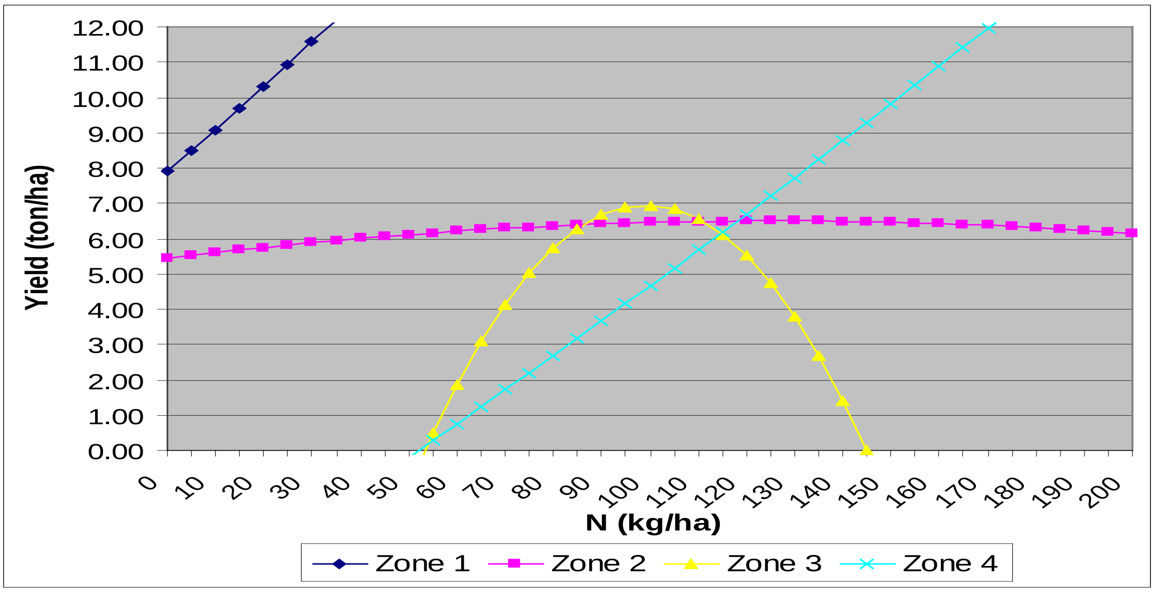

Figure 5.3: Crop response to N according to zone – Year 3 (2004/2005)

In Figure 5.3, the crop response to N according to zone for year three differs from that of the first and second years. Zones 1 and 4 shoot up in a straight line while Zone 2 is sort of linear and flat.

It can be observed from Figures 5.1 to 5.3 that the response is different for each zone and year, and that yield varies according to zone. In all the three years, Zone 3 is the only zone that displays the normal agricultural production function.

It can be concluded that, if a farmer knows the potential of different areas of the field, they can be treated accordingly, thus increasing the effective use of inputs, which can increase farming profit. This can also contribute to the reduction of production and financial risk.

Economic analysis using the partial budget is outlined in the following section. The analysis uses the coefficients estimated with the SER model for individual years to determine the profitability of VR application technology.

5.3 PROFITABILITY ANALYSIS

Profitability analysis is conducted in three phases. The first phase entails estimating profit on the basis of the average observed data, and the Baseline Model is used to estimate profit in the second phase. The sensitivity tests are used in the third and final phase to determine the optimality of the N rates used for VR treatment, as well as how they influence profitability. To conclude the analysis, a comparison is made between profit estimated by the OLS and the SER models. Because the SER model is more accurate than the OLS model, profitability is determined on the basis of the SER models. However, a comparison of the precision obtained with the SER model and with the OLS model is presented, and an analysis is done to evaluate the implications of using an inappropriate model, the OLS, which does not account for spatial effects.

5.3.1 Profit analysis: Average observed data

Farm profit is generally defined by the Department of Agriculture (2005) as remuneration to own land, capital and management. Profit, as used in this research, implies profit before remuneration of land and management (Le Clus *et al.*, 2004). It is calculated as the gross margin of the maize enterprise less fixed costs (Table 5.7 of Annexure 5) and VRT costs for the VR treatment.

As a way to conduct a preliminary profitability analysis, the average observed harvested yield and the N rates applied at the trials are used to compare profit between the two treatments. This kind of analysis is often used by farmers to evaluate the profitability of precision agriculture and to inform decision-making. The outcome is compared to the results obtained by using the coefficient estimates to determine profit. Figure 5.4 provides an overview of the profit estimates based on the actual data.

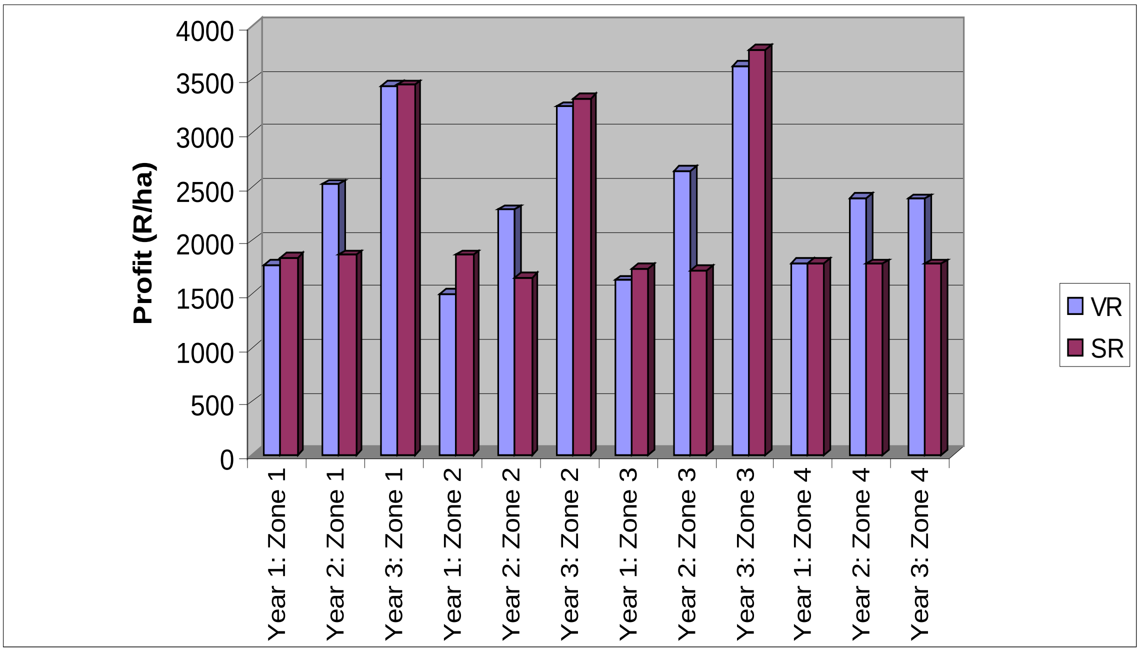

Figure 5.4: Profit estimates based on actual data

According to Figure 5.4, the highest profit is obtained from Zone 3 in the last year (2004/2005), followed by Zones 1 and 2 in the same year. Profit for the SR treatment is slightly higher in Zones 2 and 3. With the exception of these zones in the last year, as well as Zone 2 and Zone 3 in the first year, profit for the VR treatment is higher than profit for the SR in all the zones. When looking at each zone, the profit obtained under the SR treatment does not vary significantly between the zones, while an increase in profit is observed from the first to the third year for VR in all the zones. In the last year, profit stabilises for Zone 4.

Although the above analysis can provide a synopsis of the profitability of VR application technology, the coefficients estimated from the spatial regression models should form the basis of the economic analysis for profit optimisation.

5.3.2 Profit analysis: The Baseline Model

The Baseline Model is used to estimate the profitability of precision agriculture as a package. The partial budget is a tool used in the economic analysis.

The aim of most producers or farmers is to maximise profit, and it is pertinent that either the profit-maximising output or input levels are determined, since profit can be maximised from the standpoint of both the inputs used and the output produced, which are counterparts. On the output stance, profit is maximised when output corresponds to the point where the price of the output Y is equal to the marginal cost of Y ($P_y = MC_y$). A comparison is made in Figure 5.5 regarding the profit-maximising yield levels for the two treatments.

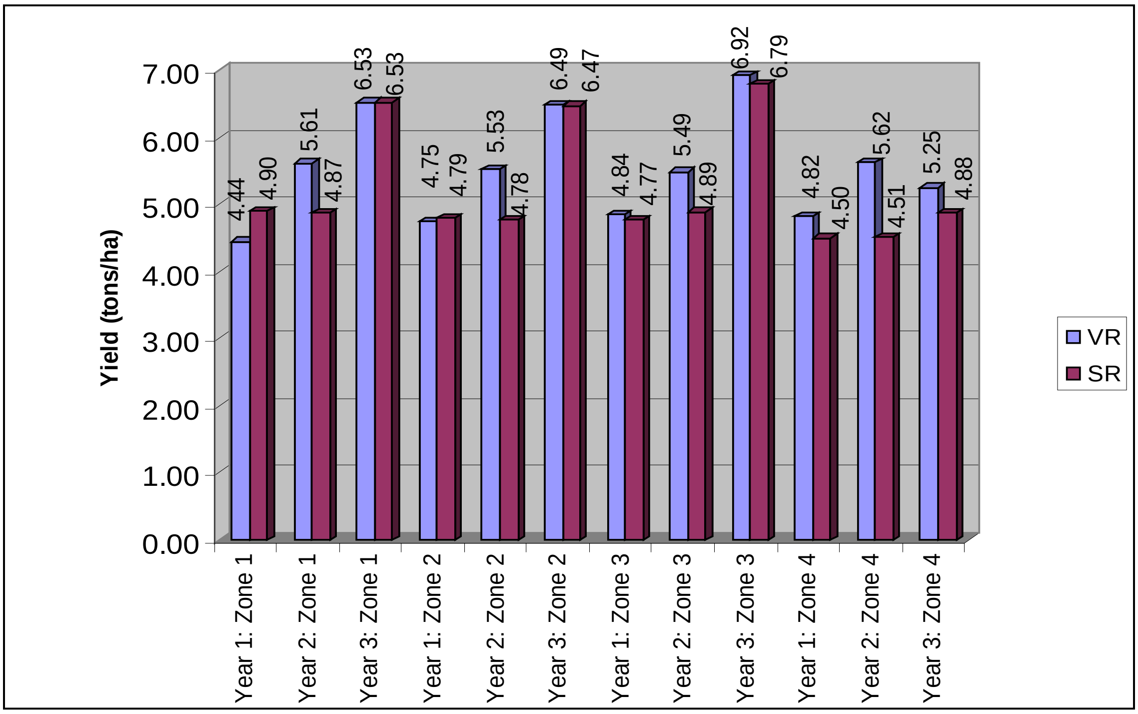

Figure 5.5: Comparison of expected yields with profit-maximizing N: Baseline Model

For Zones 1 and 2, the expected yields are higher for the SR treatment in year one, while yields are higher for VR in Zones 3 and 4. In year two, VR performs better in all the zones, whereas only Zone 4 produces substantially better yields in the third year.

Although the profit-maximising N and output levels for the two treatments should be included in the profit calculations, the average rates applied at the trials were used to calculate the profit for the Baseline Model. Separate calculations were performed for VR and SR application treatments by using formulas (1) and (2) respectively:

$$\pi = P\overline{Y}_{VR} - A - r\overline{X}_{VR} - O \qquad (1)$$

Formula 1 calculates profit for the VR application treatment, where π symbolizes profit; *P* is the price of maize yield per ton; *r* represents the price of N fertilizer in kg/ha, *O* represents other fixed and variable costs incurred in the production of maize and *A* stands for fixed costs associated with the VR application, and is calculated as follows:

$$A = I * i / [1 - (1+i)^{-n}]$$

where *I* is the investment cost of VR equipment and *i* the discount rate.

In this study, the investment in precision agriculture entails a GPS survey of the field, harnessing the tractor to put the necessary wiring in place for an existing planter, and purchasing the VR applicator, a GPS satellite receiver, a computer display, the necessary software and grid soil sampling, all totalling R320 000 (2005 prices). Training costs and the time spent by the farmer learning how to use the equipment and analysing data – although essential – are not accounted for. Additional skills and knowledge required are important with a view to estimating the economic returns. With an estimated lifetime of eight years and a discount rate of 7%[2], the annual VR costs amount to R12 673, or R121.85 per hectare.

Profit for the SR treatment is computed by formula 2, as follows:

[2] The farmer used own funds, and the 7% represents the farmer's opportunity cost.

$$\pi = P\overline{Y}_{SR} - r\overline{X}_{SR} - O \qquad (2)$$

X in both formulas is calculated with the following formula:

$$\overline{X} = \sum ij\, X_{ij} / ha$$

where X is the input applied (N), and i and j apply for VR only.

The calculations are based on a maize price of R1 000.66 per ton and N price of R2.03 per kg, both of which are three-year averages. The other production costs are based on the farmer's costs aligned with enterprise budgets for the study area (Bothaville region).

Comparison of the profit obtained from the two strategies indicates which is more profitable. For the VR, it is essential to determine whether the *possible* benefits from yield increase plus input saving are greater than the quasi-fixed costs of VR application equipment and intensive data collection (Lowenberg-DeBoer & Boehlje, 1996). As Figure 5.6 indicates, estimated profit is higher for the VR in comparison to SR in some zones.

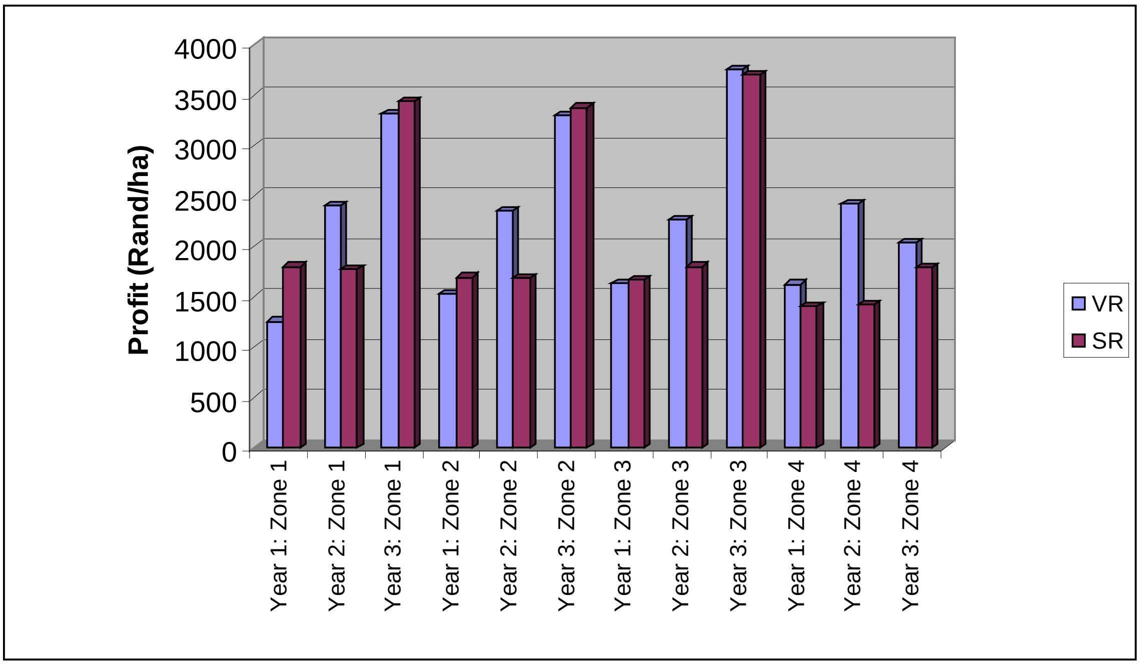

Figure 5.6: Estimated profit: Baseline Model

With the exception of Zone 4, higher profit is obtained with the SR treatment relative to the VR in years one and three. In year two, estimated profit is higher for VR in all the zones. In Zone 4, VR performs better than SR in all the years. On average, profit is R194.59 per hectare more for VR.

5.3.3 Profit analysis: Sensitivity Tests

Kahabka *et al.* (2004) are of view that the practical application of precision agriculture requires not only the determination of agronomic optimum fertilizer rates, but the economic optimum as well. Optimum fertilization levels contribute more significantly to profit than yield. Yield increases, although statistically significant, may not justify the additional expenses of the inputs (fertilizer, equipment and management) required to obtain the desired results (Kahabka *et al.*, 2004).

5.3.3.1 Profit-maximizing input levels

The yield-maximising input levels for each zone were estimated earlier in the chapter (Table 5.11), but in order to assess the profitability of the technology, income and costs must be factored in to determine the most profitable input or production level. There is a vast difference between input application for maximum production and input application for maximum profit. The latter is of essential importance in this chapter. The optimum input application rates depend on input and output price relationships. If the rate of an increase in output prices is faster than the rate on input price increase, the most profitable production and input level will be near the level associated with maximum production. However, the agricultural sector in South Africa is characterised by a cost-price squeeze, and farmers will be forced to determine optimal input levels.

From the input perspective, profit is maximised when each input is used to the point where the price of the input x is equal to the value of the marginal product ($P_r = VMP_y$), or the marginal factor cost is equal to the marginal value product (MFC = MVP). The profit-maximising levels of N for the VR treatment were determined for each zone in each of the three years, using the coefficient estimates for single-year models. Table 5.11 shows the yield-maximising rates for each zone without differentiating between the treatments, while

Figure 5.7 illustrates the optimum VR N rates for each zone in each year, estimated from the single-year Sensitivity Test 3 Model (the N model with the treatment variable). The profit-maximising N rates under the VR treatment were obtained by solving for *x* in the profit calculations of the VR treatment. For the SR, the constant rates applied at the trials were used in estimating profit.

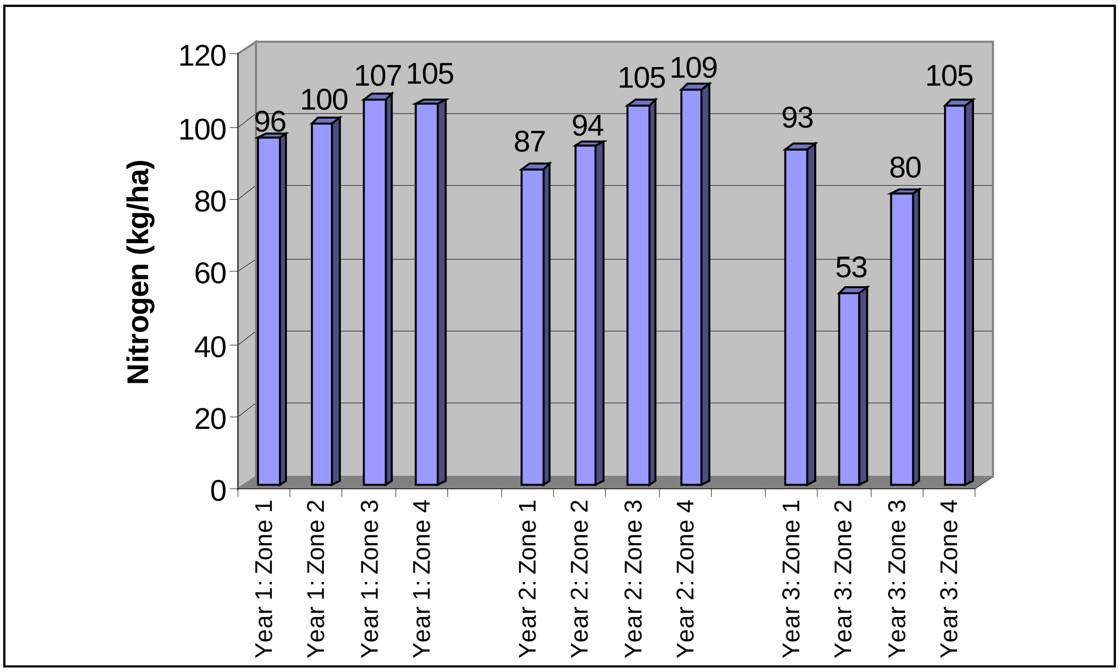

Figure 5.7: Optimal VR N rates according to year and zone (kg/ha)

The highest optimal rate is found in Zone 4 (109 kg N/ha) in year two, and the lowest optimal N (53 kg/ha) in Zone 2 in year three. The optimal N rates from the OLS do not vary significantly from those obtained with the SER model for this specification. It is important to note that profit-maximising N rates in Figure 5.7 are higher than yield-maximising rates in Table 5.11. Rates in Figure 5.7 are computed from the Sensitivity Test 3 Model, while Table 5.11 uses the Sensitivity Test 2 Model.

5.3.3.2 Profit-maximizing yield levels

In estimating the profit-maximising output levels using the sensitivity tests, the first-order necessary condition for a maximum, $d\pi/dY = 0$, had to be satisfied. This was obtained by differentiating $\pi(Y) = R(Y) - C(Y)$ with respect to Y, where Y is the output or maize yield, R is the revenue and C the cost of maize production. According to Chiang (1984), the optimum output must satisfy the equation $R'(Y)=C'(Y)$ or MR=MC, the first-order condition for maximum profit. This first-order condition can, however, result in a minimum rather than a maximum profit, and the second-order condition must also be satisfied. The first derivative provides information about the slope, while the second derivative elucidates the curvature (Chiang, 1984; Alivelu *et al.*, 2003). The second derivative can also show declining marginal products, indicating that production is taking place in the second stage, which is the optimal stage. An agricultural production function is normally represented by a concave curve with a positive first derivative and negative second derivative within a certain range of interest (Dillon & Hardaker, 1980). To satisfy the second-order condition, the first derivative was differentiated for an output level Y, in such a way that $R'(Y) = C'(Y)$. This is represented by:

$$\frac{d^2\pi}{dY^2} = R''(Y) - C''(Y),$$

which should be <0 if $R''(Y)<C''(Y)$.

This second-order condition is satisfied at an output level of Y, which results in $R''(Y)<C''(Y)$. With these conditions satisfied, the implication is that the value of the choice variable, Y, was chosen such that π will be at a maximum. Thus, a profit-maximising output was established. In economic terms this means that, if the rate of change of MR is less than the rate of change of MC at the output where MR=MC, then such an output level will maximise profit (Chiang, 1984; Hurley *et al.*, 2003).

On the basis of the single-year regression results obtained, the yield level that will maximise profit was not computed in the first derivative since - as explained earlier – the first derivative identifies the extreme values, which can be the minimum or the maximum. It was essential to determine the second derivative of yield with respect to N in order to establish whether the estimated yield is indeed the maximum. As the first-order and second-order conditions were satisfied, it can be concluded that the optimum yield was estimated. The optimal maize yields per zone for each year are estimated from the optimal N rates, and these are illustrated in Figure 5.8.

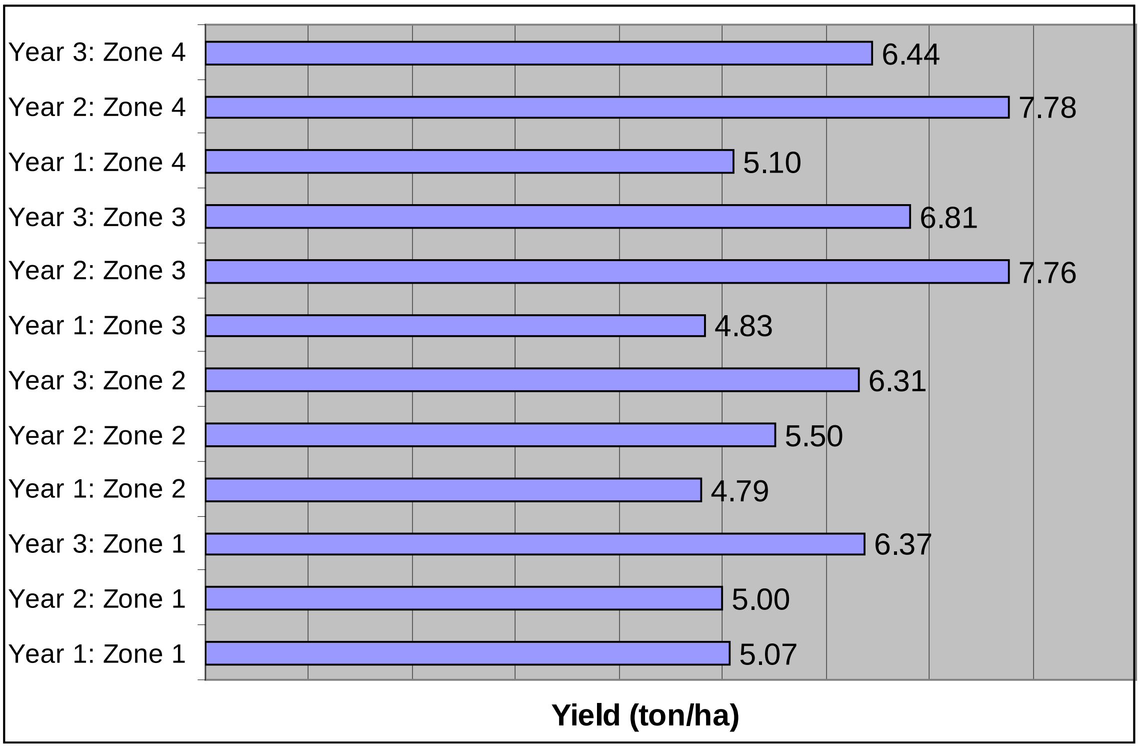

Figure 5.8: Optimal maize yields per zone and year as a response to N

In the three years of study, the higher optimal yields were obtained in Zones 3 and 4, the zones of high and very high potential, respectively. This is consistent with the identification of management zones, implying that higher yields are produced in high-potential management zones.

For a comparative analysis between the two treatments, profit-maximising output levels (yields) for different zones in different years under the two treatments were also estimated (Figure 5.9). Figure 5.8 uses Sensitivity Test 2 Model, while Figure 5.9 is generated from the results of Sensitivity Test 3 Model.

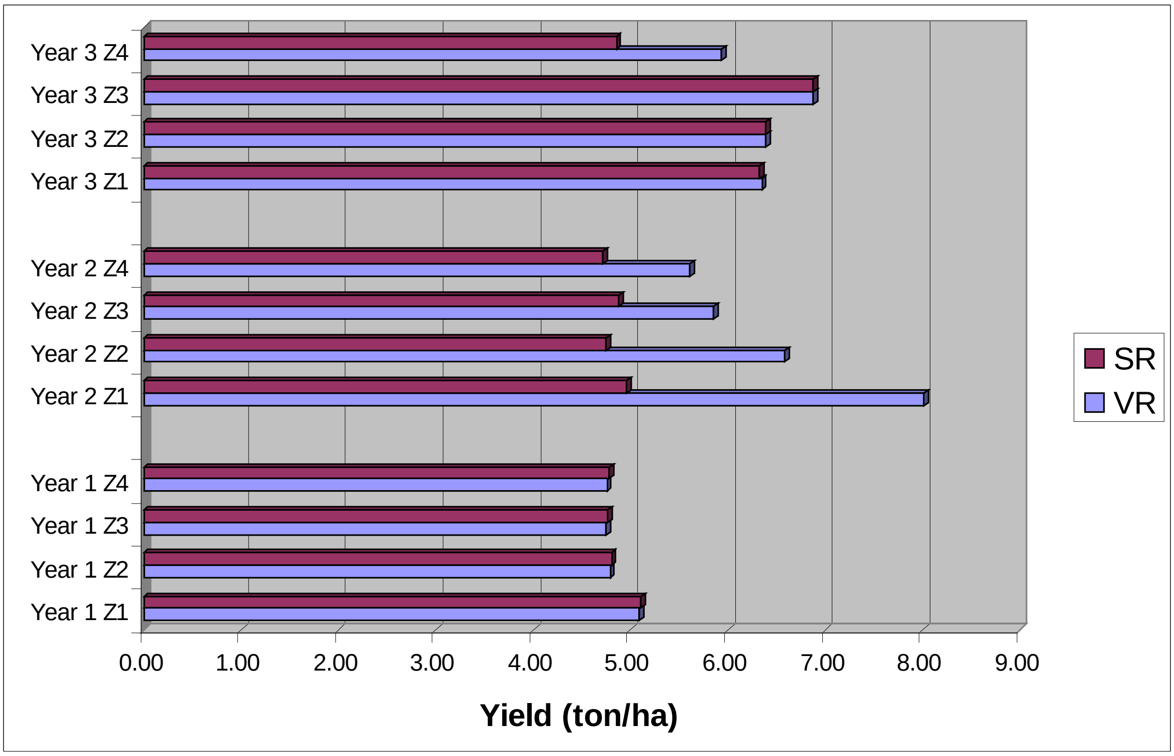

Figure 5.9: Estimated optimum yields per treatment

From Figure 5.9, it can be observed that the profit-maximising output required for the VR treatment outstrips the output required to maximise profit in the SR treatment in year two, and this is consistent with the results of the Baseline Model. As for the first and third years, the optimum yields computed for both treatments are identical. In estimating profit, costs will therefore play an important role in distinguishing the profits generated with the two treatments. In comparing the actual field yield averages with the estimated profit-maximising yield levels a major difference is found in Zone 4, with as much as 2.5 tons difference between the two in the last year, levels being higher for the latter. Such a big difference is also recorded for Zone 3 in year two. In the rest of the zones the difference is small.

5.3.3.3 Estimated profits

The profit estimate results for individual years are consistent with the results of the Baseline Model. The estimated profit for each zone, computed from the estimated optimum yields and the optimal N (for VR), is presented in Figure 5.10.

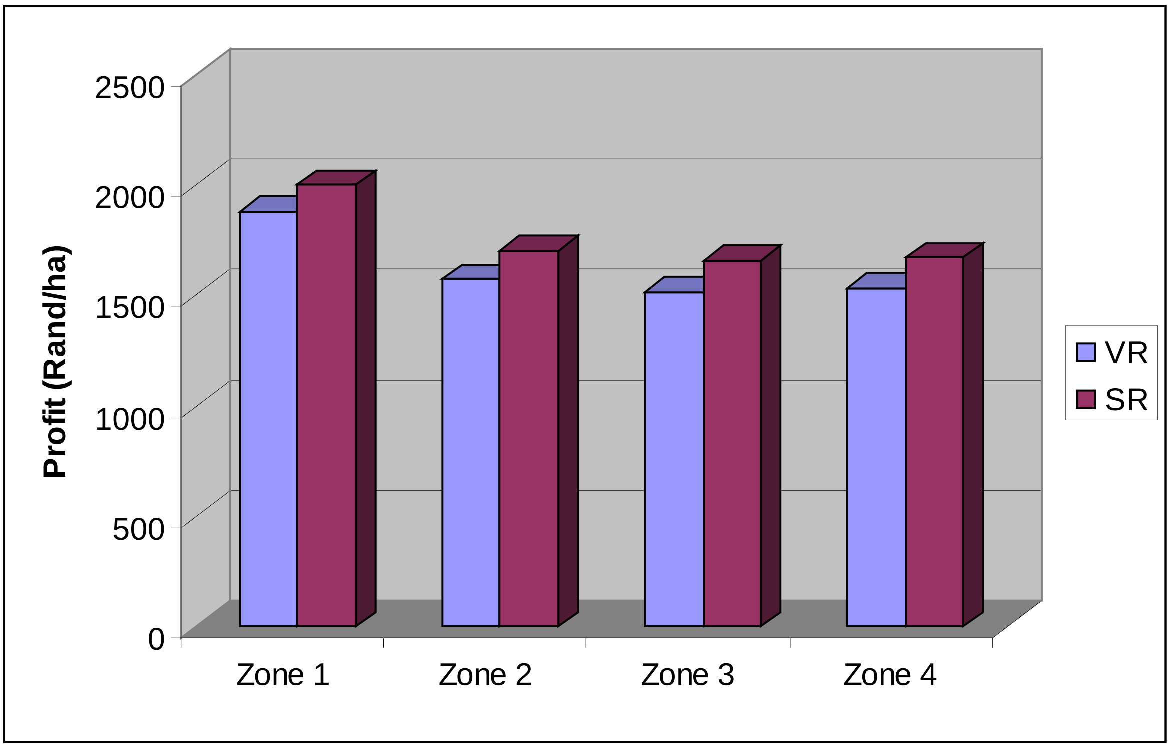

Figure 5.10: Estimated profit per zone for Year 1, using the Sensitivity Test 3 model

In the first year of study, higher profit is obtained in the SR treatment for all management zones in comparison with the profit obtained under the VR treatment, as Figure 5.10 indicates.

Although the VR variable is statistically significant at 1% level (Table 5.13), VR treatment incurs additional costs, and it is logical that a higher profit will be obtained with the SR treatment. When analysing profit per management zone in year one, the low-potential zone (Zone 1) generated the highest profit of R2 029/ha. The low-potential zone received the lowest amount of N, but the resulting yields were higher than average. Profit obtained in the other three zones for the two treatments did not vary much. Although on the whole the SR treatment performed better than VR, the difference in profit is not substantial, with an average difference of R131/ha, or R13 100 for a 100-hectare field.

As was the case with the Baseline Model, a different picture emerges in the second year. Profit estimates for the second year are shown in Figure 5.11.

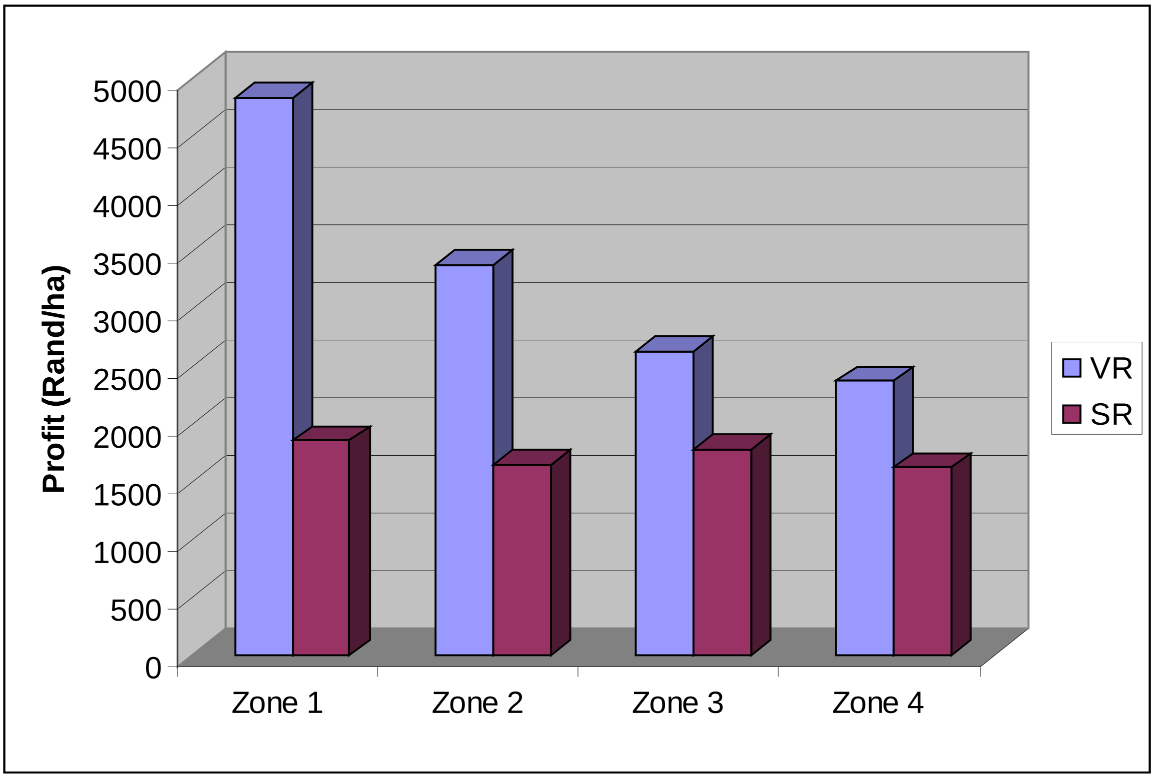

Figure 5.11: Estimated profit per zone for Year 2, using the Sensitivity Test 3 model

VR treatments resulted in the highest profit for all the zones. In comparison with the first year, profit difference is considerable, averaging R1 581/ha. Profit per zone seems not to vary between treatments only, but also across the zones as well for VR. Zone 1 has the highest profit, followed by Zones 2, 3 and 4 respectively.

Figure 5.12 demonstrates estimated profit with regard to the third year.

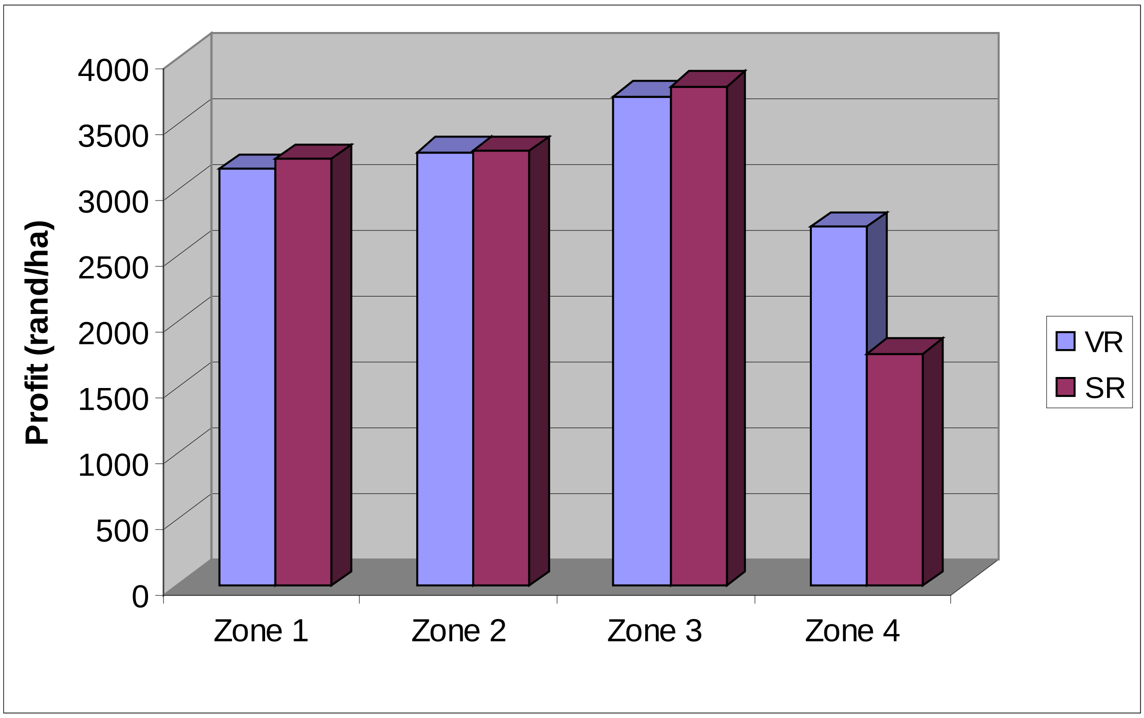

Figure 5.12: Estimated profit per zone for Year 3, using the Sensitivity Test 3 model

In the last year, with the exception of Zone 4, the two treatments resulted in profits that do not differ much. As with the Baseline Model, VR out-performed SR, mainly in Zone 4. The average difference in profit between the two treatments is R184, with higher profits for VR. Unlike the first two years, Zone 1 is not the best performer, with profit from Zone 3 superseding Zone 1 profit. The overall estimated profit per zone for the three years is collated in Figure 5.13.

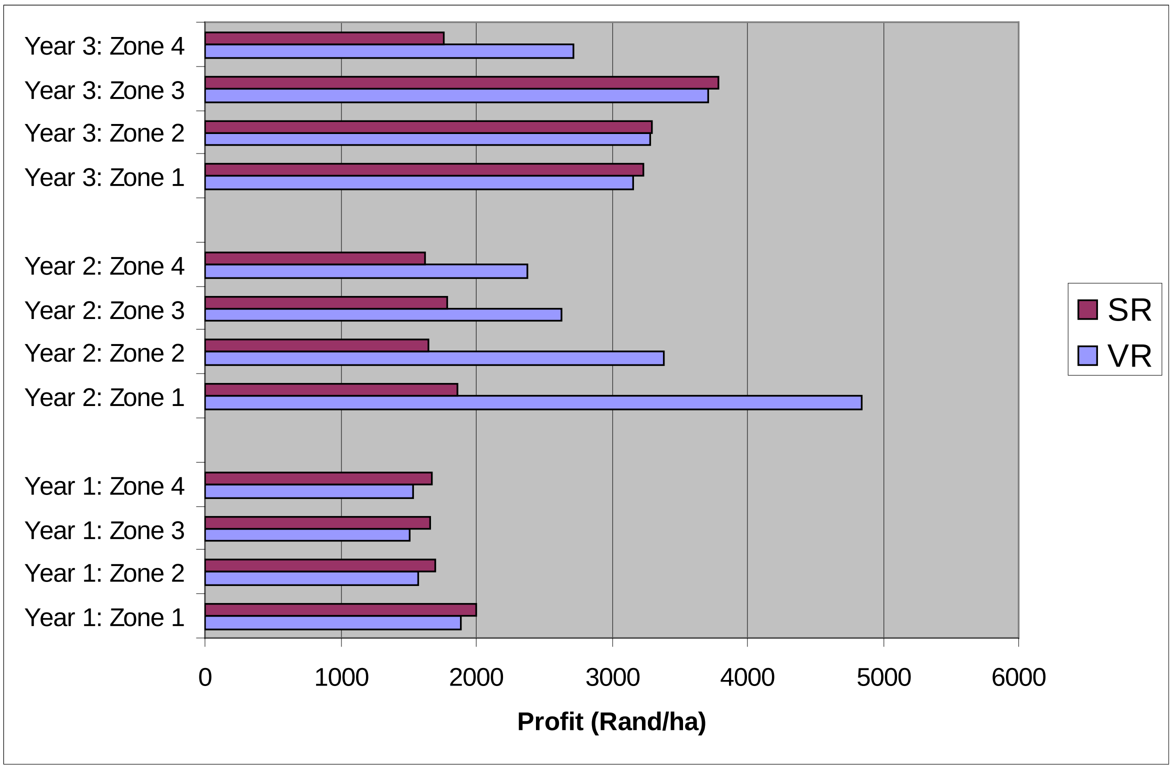

Figure 5.13: Estimated profit for the three years, using the Sensitivity Test 3 model

Although the magnitudes of the estimated profits differ, the same general conclusion is reached by determining profit estimates using the average observed yield and N application rates at the trials (Figure 5.4) and the Baseline Model (Figure 5.6), as well as the coefficient estimates of sensitivity tests. The profits are higher for the VR treatments relative to the SR treatments.

To summarize, profit estimates computed from the optimal inputs and profit maximizing output levels using the coefficient estimates of the SER models for individual years are illustrated in Table 5.15. For the SR, the constant N rates applied at the trials are used.

TABLE 5.15: PROFIT SUMMARIES FOR THE THREE YEARS OF STUDY USING THE SENSITIVITY TEST 3 MODEL

2002/2003	*Y* (ton/ha)*	*X* (kg/ha)*	*Profit (R/ha)*	**2003/2004**	*Y* (ton/ha)*	*X* (kg/ha)*	*Profit (R/ha)*	**2004/2005**	*Y* (ton/ha)*	*X* (kg/ha)*	*Profit* (R/ha)*
VR Zone 1 (Y1)	5.11	96	1 884	VR Zone 1 (Y2)	8.02	87	4 838	VR Zone 1 (Y3)	6.35	93	3 154
VR Zone 2 (Y1)	4.81	100	1 576	VR Zone 11 (Y2)	6.59	94	3 388	VR Zone 2 (Y3)	6.39	53	3 287
VR Zone 3 (Y1)	4.77	107	1 516	VR Zone 3 (Y2)	5.86	105	2 634	VR Zone 3 (Y3)	6.88	80	3 708
VR Zone 4 (Y1)	4.78	105	1 533	VR Zone 4 (Y2)	5.61	109	2 376	VR Zone 4 (Y3)	5.94	105	2 714
SR Zone 1(Y1)	5.09	105	2 002	SR Zone 1 (Y2)	4.96	102	1 849	SR Zone 1 (Y3)	6.33	102	3 227
SR Zone 2 (Y1)	4.79	105	1 703	SR Zone 2 (Y2)	4.75	102	1 650	SR Zone 2 (Y3)	6.39	102	3 294
SR Zone 3 (Y1)	4.75	105	1 658	SR Zone 3 (Y2)	4.88	102	1 780	SR Zone 3 (Y3)	6.88	102	3 782
SR Zone 4 (Y1)	5.09	105	1 672	SR Zone 4 (Y2)	4.72	102	1 624	SR Zone 4 (Y3)	4.86	102	1 759

Y*: Estimated profit-maximizing output

X*: Estimated optimum input

The profits for each zone are estimated to distinguish between performance in different zones in the three study years. Table 5.16 shows the three-year weighted and unweighted average profits.

TABLE 5.16: THREE-YEAR AVERAGE PROFITS

	SR		
	Unweighted profit	Weights	Weighted profit
Zone 1	7 089	0.21	570
Zone 2	6 647	0.49	1 491
Zone 3	7 221	0.22	3 532
Zone 4	5 055	0.08	1 040
	6 503		***6 635***
	VR		
Zone 1	9 877	0.20	843
Zone 2	8 251	0.49	1 900
Zone 3	7 858	0.22	3 229
Zone 4	6 623	0.09	1 306
	8 152		***7 877***

Over a three-year period, the estimated profit is higher for the VR treatment. In VR, Zone 1 results in the highest profit, followed by Zone 2. However, with the weighted averages, profit is higher in Zone 3, and Zone 1 has the lowest profit. For SR, the highest profit is obtained in Zone 3 – both weighted and unweighted.

5.3.4 The implications of not using the "best" model

The OLS model is inefficient in the presence of spatial autocorrelation, and results in larger coefficient estimates that may lead to inaccurate profit estimations. The purpose of this subsection is to determine the implications of using the OLS model, i.e. what conclusions could be reached when using the OLS in comparison to using the spatial regression model. A comparison of the profits estimated from the OLS and the SER models is illustrated in Table 5.17.

TABLE 5.17: COMPARISON BETWEEN PROFIT ESTIMATES USING THE SER AND OLS MODELS

	Profit - OLS	Profit - SER	Difference
		VR	
Year 1	3 033.19	1 884.44	1 148.75
	2 267.29	1 575.74	691.56
	1 858.39	1 515.51	342.87
	1 666.20	1 532.87	133.33
		VR	
Year 2	4 923.87	4 838.27	85.61
	3 473.94	3 387.65	86.29
	2 676.49	2 633.63	42.86
	2 388.01	2 376.05	11.95
		VR	
Year 3	4 877.55	3 154.11	1 723.43
	3 884.87	3 287.45	597.42
	3 755.82	3 708.47	47.35
	1 951.53	2 714.09	-762.55

The OLS model overestimates profit for VR treatment in all the study years. This implies that an inaccurate conclusion could be reached regarding the profitability of precision agriculture when using the OLS model. It could be concluded that the technology is profitable while, in fact, it may be not profitable under the specific conditions.

5.4 CONCLUSION

The results of this study indicate that there is a significant variation in maize yield response to the applied N, depending on the application method used (Table 5.6). Profit analysis based on the application strategies indicates that, on average, VR resulted in higher profits than SR (Table 5.16). The analysis indicates that, over time, the yield obtained with the VR treatment can compensate for the additional costs incurred with VR. This finding is consistent in all the models.

It has been established that yield response to fertilizer depends on soil characteristics such as the effective soil depth, which has a positive effect on yield. Yield response also differs from one management zone to the next and between the years; however, the performance of individual zones is contrary to expectations. Zone 4 was expected to perform better than the other zones; however, this was not the case due to improper identification of management zones. Differences are observed between the results obtained with the OLS model and the SER model, and this may have an impact on decision-making. This stresses the importance of taking spatial effects into account, as inaccurate results can be obtained with methodologies that ignore the spatial dependencies in the analysis of yield monitor data.

Chapter 6

SUMMARY, CONCLUSIONS AND PROPOSITIONS

6.1 INTRODUCTION

Maize response to variably band-placed nitrogen (N) was analysed over a period of three years (2001/2-2004/5) in the Bothaville district of the Free State Province in South Africa. The design and experimental procedures followed allowed investigation into the effects of different input application strategies such as variable-rate (VR) and single-rate (SR) applications on yield, without causing major disruptions in normal farming activities. The value of VR application was measured by comparing the profit from the application of N in the absence of precision information and technology, with the profit from optimal application of N with the use of precision technology and information.

This study indicates that on-farm trials, in conjunction with spatial econometrics, can be used successfully to evaluate the profitability of VR application of nitrogen. Spatial regression models were used to analyse the yield monitor and VR N application data. The usefulness of results from on-farm comparison trials can be increased when spatial autocorrelation that is present in the data is modelled correctly. Spatial models take into account the spatial effects inherent in this type of data and generate lower standard errors than Ordinary Least Squares (OLS) models. OLS models are unreliable in the presence of spatial variability because assumptions of normality, independent observations and identically and independently distributed errors are dishonoured. The use of spatial models overcame these limitations presented by OLS models. All the measures of goodness of fit indicated an increase in fit from the OLS to the spatial error (SER) model, implying that the use of these models resulted in more accurate estimates.

The study also illustrates the usefulness of collaboration between researchers, farmers and the business sector with a view to conducting new research and addressing the profitability issues of precision agriculture under local conditions. It is further demonstrated in this study that, by using on-farm strip trials, precision agriculture technologies such as VR application can be employed to assess and demonstrate alternative fertilization and other management practices made possible by the latest technology to improve the sustainability of agriculture.

The results of this study indicate that permanent soil characteristics such as the effective soil depth and the clay percentage in the A-horizon do have an effect on yield. The exploratory data analysis revealed a positive relationship between yield and these variables up to a certain point. The clay content throughout the field is relatively homogenous. The mean clay percentages between different zones do not vary statistically in the three years of study. The mean clay percentages are also not statistically different between the two treatments in all the years of study. The main distinguishing variable on the study field is the effective soil depth. Statistical differences were found on the mean effective soil depths for the zones.

The following conclusions were generated in hypothesis testing:

- There is evidence of spatial autocorrelation and heteroscedasticity in the maize response to N on this particular field.
- Spatial models result in more accurate estimates than the OLS models.
- Maize response to nitrogen and profit-maximising levels of N vary in different management zones.
- Maize response to N and profit-maximising nitrogen levels in different management zones vary in different years.
- Over a three-year period, VR application of N leads to higher economic returns compared to SR.

6.2 SUMMARY OF REGRESSION RESULTS

Baseline model for pooled data

- In year one, the VR treatment (TRT) has a positive effect mainly on Zone 4.
- The VR treatment effect on Zones 1, 2 and 3 is close to zero in years one and three.
- In year two, the VR treatment effect is positive in all zones.
- VR treatment effect varies between zones and years.

Sensitivity Test 1: The TRT Effective-depth model

- TRT has a positive effect on soils above average depth in year one.
- In year two, TRT is positive throughout the field up to a depth of 205 cm.
- TRT has a positive effect on shallow soils in year three.

Sensitivity Test 2: The Nitrogen-Zone model

- Nitrogen response varies according to zone.
- The estimated model is convex for N in year one.
- The estimated model is concave for N in years two and three.
- Yield-maximising N levels vary for different zones.
- With the exception of Zone 4, actual VR N application rates at the trial approximate yield-maximising N levels.

Sensitivity Test 3: The Nitrogen-Zone Treatment model

- N response is convex for all three study years.
- N response varies by zone.
- TRT is significantly different from zero (likelihood ratio test).

Baseline model for individual years

- There is a significant difference between yields obtained for VR and SR; VR treatment performed better than SR in all the years of study.
- In year two, VR treatment had a considerable positive effect on yield in each zone.

6.3 SUMMARY OF PROFITABILITY ANALYSIS

- Higher profit is realised for SR treatment for all management zones compared to VR in year one; however, the difference is not substantial, averaging R131/ha.
- In year two, the highest profit for all zones is obtained in the VR treatment, with a considerable difference of an average of R1 581 per hectare.
- Profits for both treatments are very close in year three, with VR performing better than SR in Zone 4 only. The average difference in profit between the two treatments is R184, with higher profits for VR.
- Over a three-year period, the estimated weighted profit is higher for the VR treatment.

6.4 THE IMPLICATIONS OF USING A WRONG MODEL

- The OLS models overestimate profits in all the study years.
- A wrong conclusion could be reached with OLS, i.e. VR application could be found to be highly profitable while, in fact, it may not be profitable in the specific case.

6.5 LESSONS LEARNT

A number of factors need to be taken into account in future research of precision agriculture. The strip-plot design is not suitable for sandy soils because of the lateral movement of water and nutrients, with spill-over effects to neighbouring strips. This is the case especially for N and potassium (K), which are readily mobile. It is believed that this might have confounded the results in the first two years of the study, even though the observations in the middle of the strips were selected for analysis. If the strip-plot design is used, it is best to leave a strip between the treatments to reduce the edge effects. This strategy was adopted in the third year of study, but the same rows were kept for each treatment. Large block trials can solve some of the problems associated with strip-plot designs. The management of treatment edge effects is efficient, as questionable points can be removed without compromising data availability (Griffin *et al.*, 2004b)

Historical yield monitor data spanning three years was used as the main base for determining management zones, but other factors – especially the permanent soil characteristics such as depth and clay – are much more important. It is not adequate to use historical yields alone to determine management zones. Before any management zones can be defined, the nutrient status of the soil, especially pH, should be brought to an acceptable level to avoid misclassification of certain areas. Some parts of the study field were classified as low-potential due to below-average yields, but it was later discovered that the low pH level was limiting the potential of those areas. It is essential that pH problems are rectified before commencing with the trials or identifying management zones.

As no systematic differences in yields between the zones were observed, it can be concluded that better zonation can be implemented to demarcate zones in a way that consistently explains yield. An improved zone map developed by Omnia, a fertilizer company in South Africa, is attached as Figure 6.1 in Annexure 6. With software developed by a fertilizer company named Omina, the five-year yield data was layered on top of soil data to produce a zone map that shows spatial trends. Figure 6.2 in Annexure 6 depicts the yield trend across the field over a five-year period, and shows some very high yields in the area initially classified as a low potential zone. This can be seen by layering a transparency of an initial hand-drawn zonation map on top of the new zonation map.

Another important lesson learnt in this study is that topography maps can play an important role in explaining variation in zone potential. The study field looks relatively flat and level to the naked eye, but the topography map revealed differences regarding drainage capabilities in various areas of the field. This can have an impact on the productivity of the soil, depending on the rainfall amount and distribution in a particular year.

A fertilizer mix was used, which complicates the separation of individual constituents of the mixture. It is better to use single fertilizers such as super-phosphates, Lime Ammonium Nitrate (LAN) and Urea Ammonium Nitrate (UAN). As these fertilizer types are more costly, a balance should be reached between the objective of the farmer, which is to reduce costs and maximise profits, and the researcher's objective, i.e. making accurate estimations.

The issue of validity should be borne in mind in precision agriculture research. It is not possible to control all extraneous factors in agriculture, as it is subjected to the influences of erratic weather and other external factors.

6.6 LIMITATIONS

Analysis of the three-year data aggregated together provided an overview of the maize crop response to N over a three-year period. However, it is not clear whether the year dummy variable is an adequate temporal model. Ignorance of temporal correlation may have the same dire results as ignorance of spatial correlation. With this limitation acknowledged, a mechanism to adequately address temporal relationships could not be found, since methods for estimating multi-year yield monitor data are not currently in place. In the analysis of this kind of data, all interactions between the variables need to be included in the model; however, some interactions were omitted due to the multicollinearity problem.

6.7 CONCLUSION

The results presented apply to field block # 7 at Rietgat farm, and responses will vary from farm to farm and even from field to field. As a result, conclusions can only be made for the field studied, and no generalisations are possible. VR application of nitrogen resulted in an increase in yield from the first to the third year on this field. On the whole, the yields were highest in the third year. The effects of on-farm comparisons also contributed to yield changes in different years. Management and knowledge increased with the years, leading to better decisions and an improvement in yields. The results obtained in the preceding years were used to adapt management practices as time went on.

Bibliography

AGR318. 1998. *Maize production*. Agronomy lecture notes. University of the Free State, Bloemfontein, South Africa.

Adamchuk, V. I., Lund, E., Dobermann, E. & Morgan, M. 2003. On-the-go mapping of soil properties using ion-selective electrodes. Paper presented at the *Fourth European Conference on Precision Agriculture*, June 15-19, 2003.

Alivelu, K., Srivastava, S., Subba Rao, A., Singh, K. N., Sevakumari, G. & Raju, N. S. 2003. Comparison of modified Mitscherlich and response plateau models for calibrating soil test based nitrogen recommendations for rice on Typic Ustropept. *Communications in Soil Science and Plant Analysis Journal*, 34(17-18).

Anselin, L. 1999. *Spatial econometrics*. University of Texas at Dallas. Texas: Richardson.

Anselin, L. 2002. *Exploring spatial data with DynESDA2*. CSISS and Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign. September 12, 2002.

Anselin, L. 2003. *GeoDa 0.9 user's guide*. Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign, IL.

Anselin, L., Bongiovanni, R. & Lowenberg-DeBoer, J. 2004. A spatial econometric approach to the economics of site-specific nitrogen management in corn production. *American Journal of Agricultural Economics*, 86(3):675-687.

Anselin, L. & Florax, R. J. 1995. *New directions in spatial econometrics*. Berlin: Springer.

Babcock, B. A. & Pautsch, G. R. 1998. Moving from uniform to variable fertilizer rates on Iowa corn: Effects on rates and returns. *Journal of Agricultural and Resource Economics*, 23(2):385-400.

Bailey, K. D. 1994. *Methods of social research*.4th ed. New York: The Free Press.

Bartlein,P. J. 2003. *Spatial Regression Analysis Lecture notes.* Department of Geography, University of Oregon. http://geography.uoregon.edu/bartlein/geog314_00/lectures/lec19.htm

Berck, P., Geoghegan, J. & Stohs, S. 1998 A strong test of the Von Liebig hypothesis. *American Journal of Agricultural Economics*, 82(4):948-955.

Berry, C. G. 1990. *An economic analysis of maize yields responses to fertilizer and rainfall variations at Dundee.* Thesis submitted for a Master of Science degree in Agricultural Economics, Department of Agricultural Economics, University of Natal, Pietermaritzburg, South Africa.

Blackmore, S. 1994. Precision farming: An introduction. *Outlook on Agriculture*, 23(4).

Boehlje, M. D. & Eidman, V. R. 1984. *Farm management*. USA: John Wiley & Sons.

Bongiovanni, R. & Lowenberg-DeBoer, J. 1999. Economics of variable rate lime in Indiana. *Proceedings of the Fourth International Conference on Precision Agriculture*, St Paul, MN ASA-CSSA-SSSA, 19-22 July 1998.

Bongiovanni, R. & Lowenberg-DeBoer, J. 2001. Precision agriculture: Economics of nitrogen management in corn using site-specific crop response estimates from a spatial regression model. *American Journal of Agricultural Economists Association Annual Meeting*, August, 2001.

Bongiovanni, R., and J. Lowenberg-DeBoer (2004). Precision Agriculture and Sustainability. *Journal of Precision Agriculture*, 5:359–387.

Boyer, J., Erickson, B. & Hawkins, S. 2005. Using site-specific tools to manage field-scale research. Site-Specific Management Centre Newsletter, March 2005. Purdue University, Indiana.

Brouder, S. & Nielsen, R. 2000. On-farm research. *In:* J. Lowenberg-DeBoer & B. Erickson (Eds). *Precision farming profitability,* Agricultural Research Programs, Purdue University, 103-112.

Bullock, D. G., Bullock, E. D., Nafziger, E. D., Doerge, T., Paszkiewics, S., Carter, P. & Peterson, T. A. 1998. *Does variable rate seeding of corn pay?* University of Illinois, Urbana, Illinois.

Burt, T. P., Heathwaite, A. L. & Trudgill, S. T. 1993. Overview-the nitrate issue. *In:* T. P. Burt, A. L. Heathwaite & S. T. Trudgill (Eds), *Nitrate: Processes, patterns and management*. England: John Wiley & Sons, Chapter 11:321-339.

Chiang, A. C. 1984. *Fundamental methods of mathematical economics.* 3rd ed. McGraw-Hill.

Clay, D. E., Malo, D. D., Reese, C. & Carlson, C. G. 2003 *Identifying good candidates for precision phosphorous management*. Site-Specific Management Guideline–13. College of Agriculture and Biological Sciences, South Dakota State University, Brookings.
http://www.ppi-far.org/ssmg

Dampney, P. M. R. & Moore, M. 1999. *Precision agriculture in England: Current practice and research-based advice to farmers.* Proceedings of the Fourth International Conference on Precision Agriculture. Madison, WI.

Deckard, E. L., Tsai, C. Y. & Tucker, T. C. 1984. Effect of nitrogen nutrition on quality of agronomic crops. *In:* R. D. Hauck (Ed.). *Nitrogen in crop production.* ASA-CSSA-SSSA. Wisconsin, USA, Chapter 41, 601-615.

Department of Agriculture. 2005. *Some agricultural economic concepts*. (2nd revised edition). Deierctorate Agricultural Information Services. Pretoria.

Department of Environmental Affairs and Tourism South Africa. 1999. *Bio-safety in South Africa*. http://www.ngo.grida/soesa/nsoer/general/about.htm (www.environment. gov.za).

Dillon, J. I. & Hardaker, J. B. 1980. *Farm management research for small farmer development.* FAO Services, Bulletin 41. Rome.

Drummond, S. 2005. *Yield editor 1.00 Beta version user's manual*. http://www.fse.missouri.edu/ars/ye/yield_editor_manual.pdf

English, C. B, Mahajanashetti, S. B. & Roberts, R. K. 1999. Economic and environmental benefits of variable rate application of nitrogen to corn fields: Role of variability. *American Agricultural Association Meeting*, August 8-11, 1999. Nashville, Tennessee.

English, C. B., Roberts, R. K. & Mahajanashetti, S. B. 2003. *Spatial break-even variability for variable rate technology adoption.* Department of Agricultural Economics, The University of Tennessee, Knoxville, Tennessee.

Erickson, B. 2004. Field experience validates on-the-go soil pH sensor. *Site-Specific Management Newsletter*, December, Purdue University.

Fertilizer Society of South Africa (FSSA). 1989. *Fertilizer handbook*. 3rd revised ed.Translated by E. B. Dickinson, Esmé Greenfield & J. H. Roos. Hennopsmeer, South Africa: Fertiliser Society of South Africa.

Forcella F. 1993. *Value of managing within-field variability: Soil specific crop management*. Wisconsin, USA: Madison.

Franzen, D. 1999. *Soil sampling and variable-rate fertilizer application. Site-Specific Farming*, Number 2. North Dakota State University, NDSU Extension Service, SF-1176 (2).

Gandonou, J., Dillon, C. R., Harman, W. & Williams, J. R. 2004. *Precision farming as a tool to reducing environmental damages in developing countries: A case study of cotton production in Benin.* Selected paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Denver, Colorado.

Goddard, T. 1997. What is precision farming? *Proceedings of Precision Farming Conference*, January, 20-21. Taber, Alberta, Canada.

Godwin, R. J., Earl, R., Taylor, J. C., Wood, G. A., Bradley, R. I., Welsh, J. P., Richards, T., Blackmore, B. S., Carver, M. J., Knight, S. & Welti, B. 2002. *'Precision farming' of cereal crops: A five-year experiment to develop management guidelines.* HGCA Project Report 267, London: Home-Grown Cereals Authority (HGCA).

Griffin, T. W., Lambert, D. M. & Lowenberg-DeBoer, J. 2004a. Testing appropriate on-farm trial designs and statistical methods for precision farming: A simulated approach. *7th International Conference of Precision Agriculture*, 25-28 July. Minneapolis, MN.

Griffin, T. W., Lowenberg-DeBoer, J., Lambert, D. M., Peone, J., Payne T. & Daberkow, S. G. 2004b. *Adoption, profitability and making better use of precision farming data*. Staff paper #04-06. Department of Agricultural Economics, Purdue University.

Griffin, T. W., Brown, J. P. & Lowenberg-DeBoer, J. 2005a. *Yield monitor data analysis: Data acquisition, management and analysis protocol*. Version 1. Site-Specific Management Centre. Department of Agricultural Economics, Purdue University.

Griffin, T. W., Fitzgerald, G., Lambert, D. M., Lowenberg-DeBoer, J., Barnes, E. M. & Roth, R. 2005b. Testing appropriate statistical methods for on-farm cotton research using precision farming. *Proceedings of the Beltwide Cotton Conferenc*e, 4-7 January, New Orleans, LA.

Griffin, T. W., Lowenberg-DeBoer, J. & Florax, 2006. Improving farm management decision-making experience from spatial analysis of yield monitor data from field scale on-farm trials. *8th International Precision Agriculture Conference*, 23-26 July, Minneapolis, MN.

Gujarati, D. N. 2003. *Basic econometrics*. 4th Ed. New York: McGraw-Hill Higher Education.

Hancock, D. 2002. *Some editorial thoughts of the PA topics from researchers at UK: Highlights of the agronomy meetings*. A Precision Agriculture Newsletter. UK Cooperative Extension Service, University of Kentucky, College of Agriculture.

Harris D. 1997. *Risk management in precision farming*. United Kingdom: Bios Scientific Publishers.

Heady, E. O. & Dillon, J. L. 1961. *Agricultural production functions*. Ames, Iowa: Iowa State University Press.

Henao, J. & Baanante, C. 1999. Nutrient depletion in the agricultural soils of Africa. *Vision 2020*, Brief no. 62. International Food Policy Research Institute (IFPRI).

Hurley, T. M., Malzer, G. L. & Kilian B. 2003. *Estimating site-specific nitrogen crop response functions: A conceptual framework and geo-statistical model*. Paper presented at the American Agricultural Economics Association Annual Meeting, Montreal, Canada.

Isik, M., Khanna, M. & Winter-Nelson, A. 1999. *Investment in site-specific crop management under uncertainty*. Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, Illinois, USA.

Jensen, T. & Hall, L. 2000. *Precision spraying towards an ideal.* Agrium and Alberta Food and Rural Development.
http://www.precisionfarming.com/features/0199feat-pf.html

Johnson, J. W. 2001. *Soil sampling, handling and testing. Agronomy facts*. The Ohio State University Extension Department of Horticulture and Crop Sciences, Columbus, Ohio.

Kahabka, J. E., Van Es, H. M., McClenahan, E. J. & Cox, W. J. 2004. Spatial analysis of maize response to N fertilizer in Central New York. *Precision Agriculture,* 5:463-476.

Kennedy, J. O. S., Whan, I. F., Jackson, R. & Dillion, J. L. 1973. Optimal fertilizer carryover and crop recycling policies for a tropical grain crop. *Australian Journal of Agricultural Economics,* 17:104-113.

Kleinjan, J., Chang, J., Wilson, J., Humburg, D., Carlson, G., Clay, D. & Long, D. 2004. *Cleaning yield data*. Plantsci.sdstate.edu/precisionfarm/paper/Cleaning%20Yield%20Data.pdf.

Lal, R. 1991. Soil structure and sustainability. *Journal of Sustainable Agriculture*, 1(4):67-92. The Haworth Press.

Lambert, D. & Lowenberg-DeBoer, J. 2000. Precision agriculture profitability review. *Site-specific Management Centre Newsletter*, September. Purdue University.

Lambert, D. M., Lowenberg-DeBoer, J. & Bongiovanni, R. 2002. *Spatial regression, an alternative statistical analysis for landscape scale trials: Case study of variable rate nitrogen application in Argentina*. Paper presented at the 6th International Conference on Precision Agriculture, Minneapolis, MN. USA. 14-17 July, 2002.

Lambert, D. M., Lowenberg-DeBoer, J. & Bongiovanni, R. 2003. *Spatial regression models for yield monitor data: A case-study from Argentina*. Paper presented at the American Agricultural Economics Association Annual Meeting, Montreal, Canada.

Lambert, D., Lowenberg-DeBoer, J. & Bongiovanni R. 2004. A comparison of four spatial regression models for yield monitor data: A case study from Argentina. *Journal of Precision Agriculture*, 5(6):579-600.

Lau, L. J. 1986. Functional forms in econometric model building. *A handbook of econometrics*, Vol. 3, Chapter 26:1515-1564.

Leedy, P. D. 2001. *Practical research: Planning and design*. New Jersey: Prentice-Hall.

Le Clus, C.F., Spies, D.C., Taljaard, P.R., Jooste, A. & Van Schalkwyk, H.D. 2004. *Methodology for the determination of the average production cost of maize*. Agrimark Trends AMT (Pty) Ltd.

Le Roux, P. A. L, Schoeman, J. L., Snyman, K. & Verster, E. 1997. *Guidelines for mapping and interpretations of the soils of South Africa.* Unpublished book. Department of Soil Science, University of the Free State, Bloemfontein, South Africa.?

Le Roux, P. A. L., Ellis, F., Merryweather, F. R., Schoeman, J. L., Snyman, K., Van Deventer, P. W. & Verster, E. 1999. *Guidelines for mapping and interpretation of South African soils*.

Unpublished pamphlet. Department of Soil Science, University of the Free State, Bloemfontein, South Africa.

LeSage, J. P. 1998. *Spatial econometrics*. MATLAB Guide. Department of Economics, University of Toledo.

Lewis, H. 2003. Determining the minimum acreage for cost-effective adoption of auto-guidance systems in cotton production. ARE Update, 6(6). *Giannini Foundation of Agricultural Economics. Agricultural and Resource Economics Update*. University of California, Davis, California.
http://repositories.cdlib.org/giannini/areu/UpdateV6N6/

Lowenberg-DeBoer, J. & Boehlje, M. D. 1996. Revolution, evolution or dead-end: Economic perspectives on precision agriculture. *Proceedings of the Third International Conference on Precision Agriculture*, 923-942. Madison, Wisconsin.

Lowenberg-DeBoer, J. & Swinton, S. M. 1997. Economics of site-specific management in agronomic crops. *In:* F. J. Pierce & E. J. Sadler (Eds). *The state of site-specific management for agriculture*. ASA-CSSA-SSSA, Winsconsin, USA. Chapter 16:369-396.

Lowenberg-DeBoer, J. 1998. Economics of variable rate planting of corn. *Proceedings of the Fourth International Conference on Precision Agriculture*, St. Paul, MN, 19-22 July, 1643-1651.

Lowenberg-DeBoer, J. 1999a. Risk management potential of precision farming technologies. *Journal of Agricultural and Applied Economics*, 31(2):275-285.

Lowenberg-DeBoer, J. 1999b. GPS based guidance systems for agriculture. *Site-Specific Management Newsletter*, December. Purdue University.

Lowenberg-DeBoer, J. 2002. Blocks vs. strips for on-farm experimentation. *Site-Specific Management Newsletter*, June. Purdue University.

Lowenberg-DeBoer, J. 2003. Soil pH sensor commercialised. *Site-Specific Management Newsletter*, December. Purdue University.

Lowenberg-DeBoer, J. 2004. The management time economics of on-the-go sensing for nitrogen application. *Site-Specific Management Newsletter*, May. Purdue University.

Ludden, T. M., Beal, S. L. & Sheiner, L. B. 1994. Comparison of the Aikaike information criteria, the Schwarz and the F-test as guides to model selection. *Pharmacokirat Biopahrm*, (5):481-445.

Mallarino, A. 1998. *Using precision agriculture to improve soil fertility management and on-farm research.* Integrated Crop Management. Department of Entomology. Iowa State University.

Malzer, G. L., Mamo, M., Mulla, D., Bell, J., Graff, T., Strock, J., Porter, P., Robert, P., Eash, N., Braum, S. & Dikici, H. 1999. Economic Benefits and risks associated with site-specific nutrient management. *Proceedings of the 5th Site-specific Farming Workshop*, Feb 25, 1999. Fargo, ND.

Manitoba Agriculture and Food, Factsheet and Publications. 2001. *An overview of precision farming technology*: *Is it for you?* Manitoba Agriculture and Food, Canada.
http://www.gov.mb.ca/agriculture/facts/fbc08s00.html

Matela, N. 2001. *The status of precision agriculture in cash crop production in South Africa*. Thesis submitted in partial fulfilment for the Master's degree in Agricultural Management, Department of Agricultural Economics, University of the Free State, South Africa.

Mausbach, M. J., Lytle, D. J. & Spivey, L. D. 1992. Application of soil survey information to soil specific farming: Soil specific crop management: Adoption patterns and incentives. *Agronomy Journal. (98): 381-389.*

Morgan, M., Parsons, S. & Ess, D. 2000. Global positioning systems. *In:* J. Lowenberg-DeBoer & B. Erickson (Eds), *Precision farming profitability*. Agricultural Research Programs, Purdue University, 56-61.

Morrison, J. & Pearce, R. 2000. Interrelationship between economic policy and agri-environmental indicators: An investigating framework with examples from South Africa. *Ecological Economics*, 34:363-37.

Mortensen, J. R. & Beattie, B. R. 2005. Does choice of response function matter in setting maximum allowable N-application rates in Danish agriculture? Research Paper 2005-01. *Cardon Research Papers in Agricultural and Resource Economics*, February 2005.

Moss, C. B. & Schmitz, T. G. 1999. *Investing in precision agriculture*. Department of Food and Resource Economics at the University of Florida, and Morrison School of Agribusiness and Resource Management at Arizona State University, respectively.
http://www.east.asu.edu/msabr/research/workingpapers/schimtzprecag9904.p

Napier, R. 2001. Creating opportunities from global challenges in agriculture. *Canadian Farm Business Management Council Excellence Conference*, Mont Tremblant, Quebec, Canada. 29 November 2001.

National Research Council (NRC). 1997. *Precision agriculture in the 21st century: Geospatial and information technologies in crop management*. Washington, D.C.: National Academy of Sciences.

Nell, W. T., Maine, N. & Basson, P. M. 2006. Africa, Part III: Current status. *In:* A. Srinivasan (Ed.), *Handbook of precision agriculture: Principles and applications*. New York: Food Products Press, Chapter 17:465-500.

Nelson, M. R., Orum, T. V., Jaime-Garcia, R. & Nadeem, A. 1999. Application of geographic information systems and geostatistics in plant disease epidemiology and management. *The American Phytopathological Society. Plant Disease*, 83(4).

NIST/SEMATECH. 2004. *E-handbook of statistical methods*.
http://www.itl.nist.gov/div898/handbook

Olson, R. A. 1984. Nitrogen use in dryland farming under semi-arid conditions. In: R. D. Hauck (Ed.), *Nitrogen in crop production*. ASA-CSSA-SSSA. Wisconsin, USA, Chapter 22:335-347.

Olson, K. 1999. *Is precision agriculture profitable? Does it pay?* Presentation given to international students at the Department of Applied Economics, Purdue University.

O'Neal, M., Frankenberger, J. R., Ess, D. R. & Lowenberg-DeBoer, J. 2004. Profitability of on-farm precipitation data for nitrogen management based on crop simulation. *Precision Agriculture*, 5:153-178.

Parkhomenko, P., Robert, P., Rogasik, J. & Schnug, E. 1999. *Profitability of precision agriculture: A case study from Germany*. Federal Agricultural Research Institute (FAL), Braunschweig, Germany and Department of Soil, Water, and Climate, University of Minnesota, USA.

Parkinson, R. J. 1993. Changes in agricultural practices. *In:* T. P. Burt, A. L. Heathwaite & S. T. Trudgill (Eds.). *Nitrate: Processes, patterns and management*. England: John Wiley & Sons, Chapter 11:321-339.

Pringle, M. J., Cook, S. E. & McBrateney, A. B. 2004. Field scale experiments for site-specific crop management. Part 1: Design considerations. *Precision Agriculture,* 5:617-624.

Rains, G. C. & Thomas, D. L. 2000. Precision farming: An introduction. *Georgia Extension Bulletin 1162*. University of Georgia, College of Agricultural and Environmental Sciences. http://www.ces.uga.edu/pubcd/B1208.htm

Rains, G. C., Thomas, D. L. & Vellidis, G. 2001. Soil-sampling issues for precision management of crop production. *Applied Engineering in Agriculture,* 17(6):769-775.

Roberson, G. T. 2000. *Precision agriculture: A comprehensive approach*. Biological and Agricultural Engineering Department, North Carolina State University.

Rogasik, J., Schroeder, D., Schnug, E., Schaak, G., Simchen, H. & Schrenk, L. 1999. *Operational precision agriculture on a large scale farm in Eastern Germany*. ASA-CSSA-SSA. Madison, Wisconsin.

Rust, L. 2005. Agricultural geo-referenced information system (AGIS). Personal communication.

Safo-Kantanka, O. 1994. The growth and development of Xanthosoma and Colocasia spp. under different spacings. *Acta-Horticulturae*, 380:473-477.

SAS Institute Inc. 1999. *SAS online Doc®, Version 8.* Cane, NC, USA.

Schilfgaarde, J. V. 1999. *Is precision agriculture sustainable?* American Journal of Alternative Agriculture. Vol. 14, No 1. (43-46).

Schnitkey, G. D., Hopkins, J. W. & Tweeten, L. G. 1996. An economic evaluation of precision fertilizer applications on corn-soybean fields. *Proceedings of the Third International Conference on Precision Agriculture*, 23-26 June. ASA-CSSA-SSSA.

Schnug, E., Panten, K. & Haneklaus, S. 1998. Sampling and nutrient recommendations – The future. *Commn. Soil Sci. Plant Anl*., 29(11-14):1455-1462. Institute of Plant Nutrition and Soil Science, Federal Agricultural Research Centre, Braunschweig, Germany.

Shatar, T. 1998. *Does precision agriculture make economic sense*? Australian Centre for Precision Agriculture.

Simpson, K. 1986. *Fertilizers and manures*. New York. Longman Group Limited.

SkyMark Corporation. 2004. http://www.skymark.com/index.asp

Soil Classification Working Group. 1991. *Soil classification – A taxonomic system for South Africa.* Department of Agricultural Development, Pretoria.

Spatial Analysis Laboratory. 2004. *GeoDa^{TM} spatial analysis software*. Department of Agricultural and Consumer EconomicsM University of Illinois, Urbana-Champaign, Illinois.

Tiffany, D. G. & Eidman, V. 1998. *Grower paths to profitable usage of precision agriculture technologies*. Department of Applied Economics, University of Minnesota.

Tenkorang, F. & Lowenberg-DeBoer, J. 2004. Observation on the economics of remote sensing in agriculture. *Site-Specific Management Newsletter*, March. Purdue University..

Trimble Navigation Limited. 2005. *All about DGPS.* http://www.trimble.com/gps/whydgps/html

Van Zyl, J., Kirsten, J. F., Coetzee, G. K. & Blignaut, C. S. 1999. *Finance and farmers: A financial guide for farmers.* Johannesburg: The Standard Bank of South Africa Ltd.

Varsa, E. C, Ebelhar, S. A., Chong, S. K., Indorante, S. J & Wyciskalla, T. D. 2000. Evaluation of variable rate technology as a management toll for potassium fertilization of grain crops. *Proceedings of Illinois Fertilizer Conference,* January 24-26. Peoria, Illinois.

Vermeulen, S. E., Steen, I. & Schnug, E. 1998. Nutrient balances at farm level. *In:* Codes of good fertilizer practice and balanced fertilization. *11th International Centre of Fertilizer (CIEC) International Symposium*, September 27-29. Pulawy, Poland.

Voortman, R. L. & Brouwer, J. 2001. *An empirical analysis of the simultaneous effects of nitrogen, phosphorous and potassium in millet production on spatially variable fields in SW Niger*. Staff Working Paper 01-04. Centre for World Food Studies, Department of Economic and Econometric, Vrije Universiteit, Amsterdam.

Wallace, M. B. & Knausenberger, W. I. 1997. *Inorganic fertilizer use in Africa: Environmental and economic dimensions*. Environmental and natural resources policy

Watson, M. & Lowenberg-DeBoer, J. 2002. Who will benefit from GPS auto guidance in the corn belt? *Site-Specific Management Newsletter*, December. Purdue University.

Weiss, M. D. 1996. Precision farming and spatial economic analysis: Research challenges and opportunity. *American Journal of Agricultural Economics,* 78:1275-1280. American Agricultural Economics Association.

Welsh, J. P., Wood, G. A., Godwin, J. C., Taylor, R., Earl, B. S., Blackmore, B. S., Spoor, G., Thomas, G. & Carver, M. 1999. Developing strategies for spatially variable nitrogen application. *2nd European Conference of Precision Agriculture*. Odense, Denmark, 11-15 July.

Wittig, T. A. & Wicks, Z. A. 1999. *Simple On-farm comparisons*. Site-Specific Management Guidelines. SSMG-18. Potash & Phosphate Institute, Norcross, GA.

http://www.terraseer.com/products/stis/help/default.htm#Statistics/LM/Results/Interpreting univariate_Local_Moran_statistics.htm

List of Annexures

ANNEXURE 4.4

ANNEXURE 5.1

ANNEXURE 6

Annexure 3

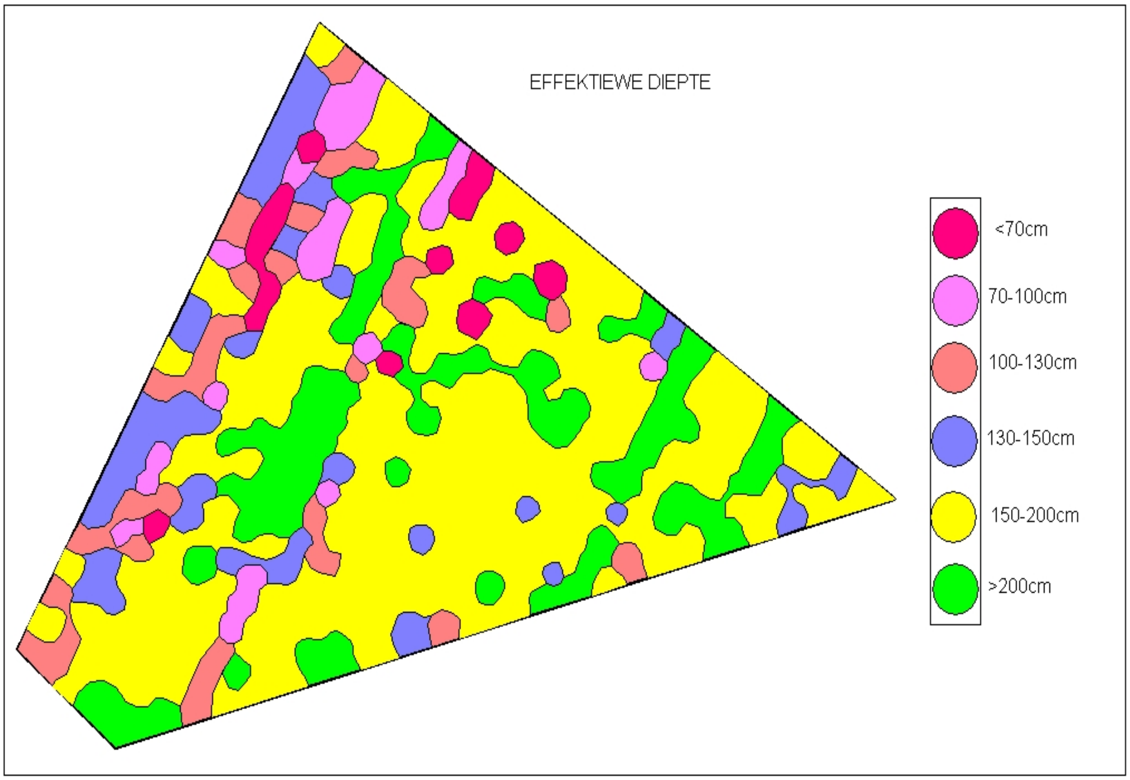

Figure 3.1: Soil effective depth over the field

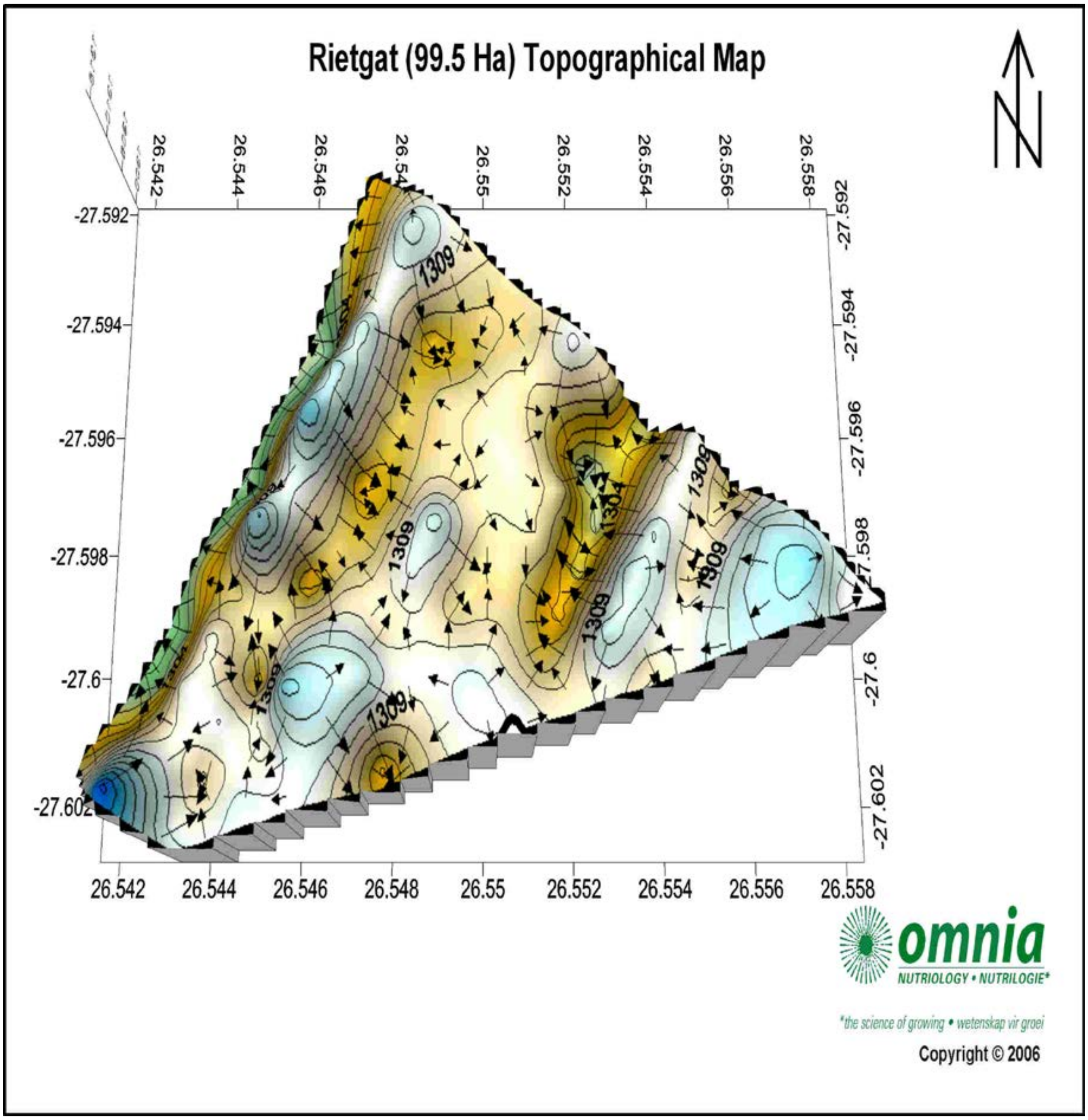

Figure 3.2: A topography map

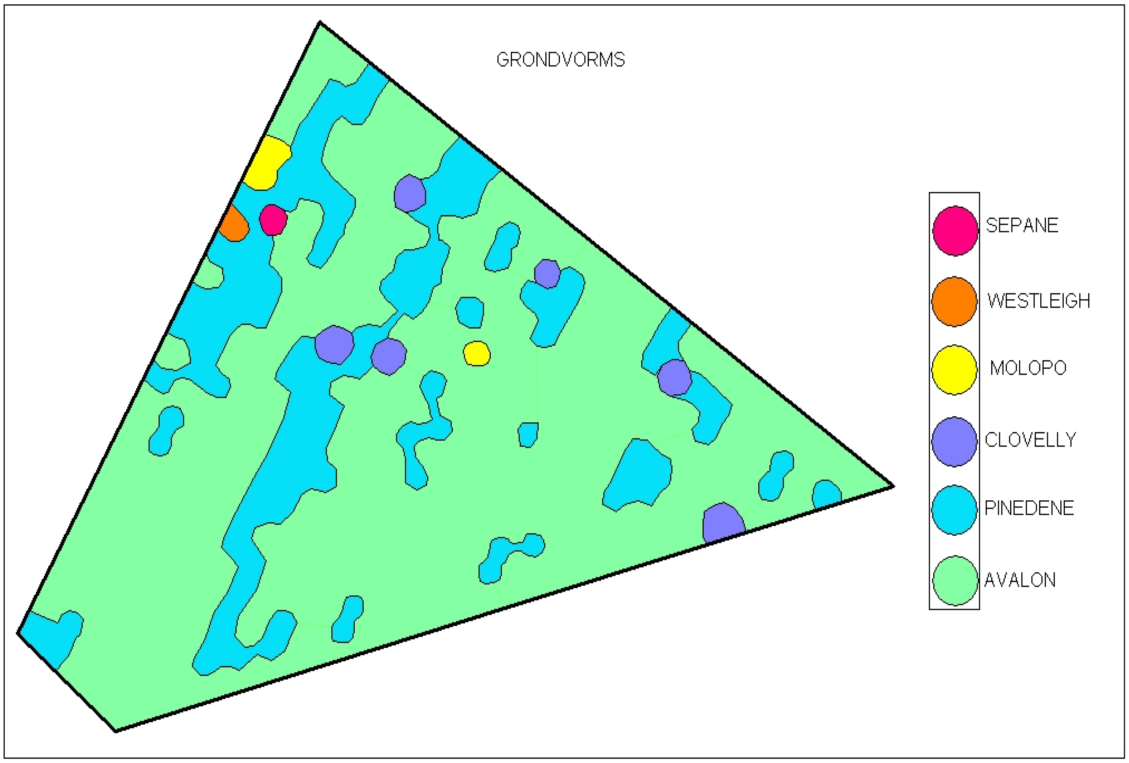

Figure 3.3: Soil forms

Table 3.1: Nitrogen application guideline

Target yield (ton/ha)	1	2	3	4	5	6	7	8	9	10
Kg Nitrogen per ha	12	36	60	84	108	134	158	184	208	232

Source: Fertilizer Society of South Africa (1989)

Annexure 4

Annexure 4.1

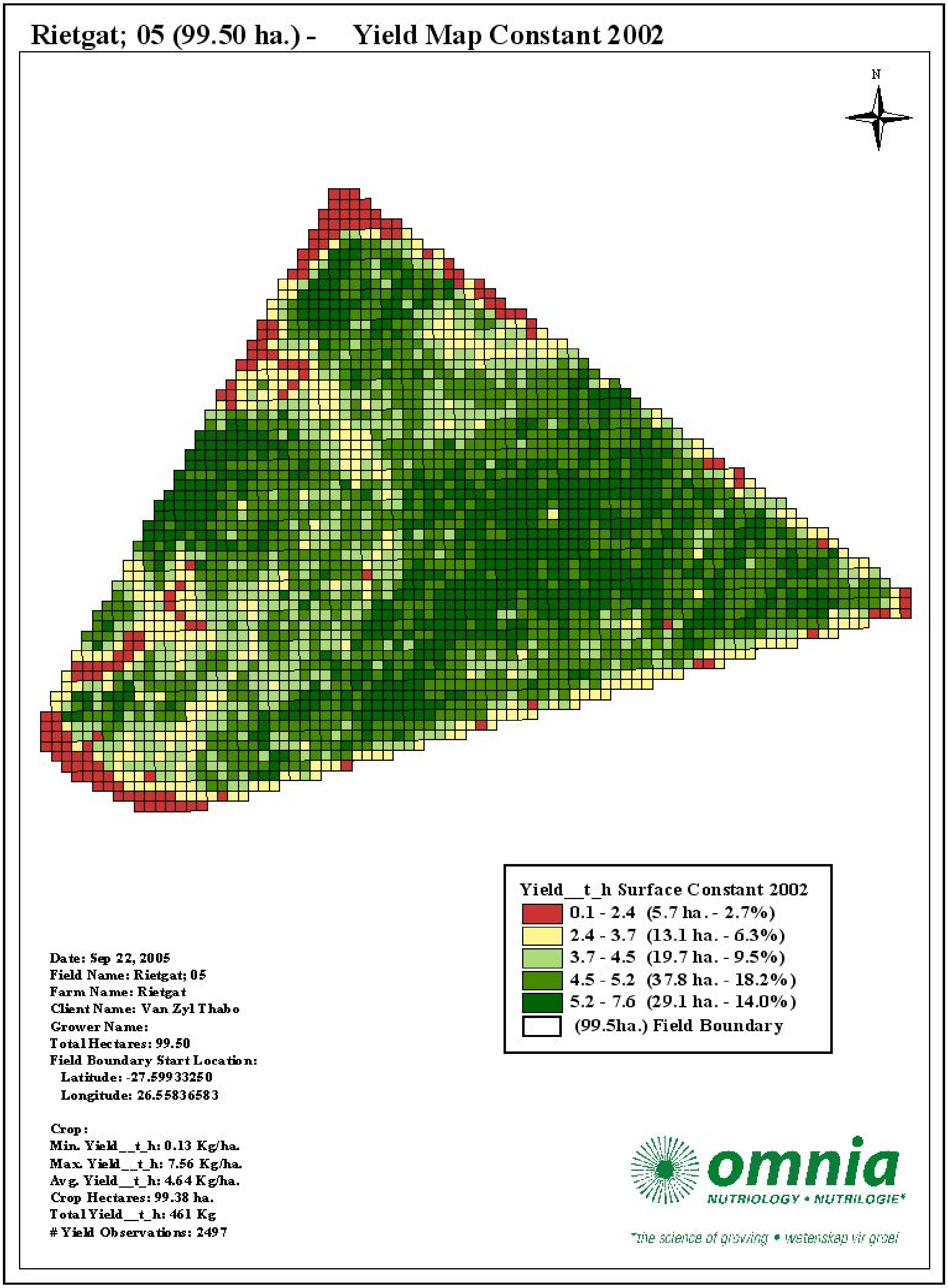

Figure 4.1a: SR Yield map 2002/2003

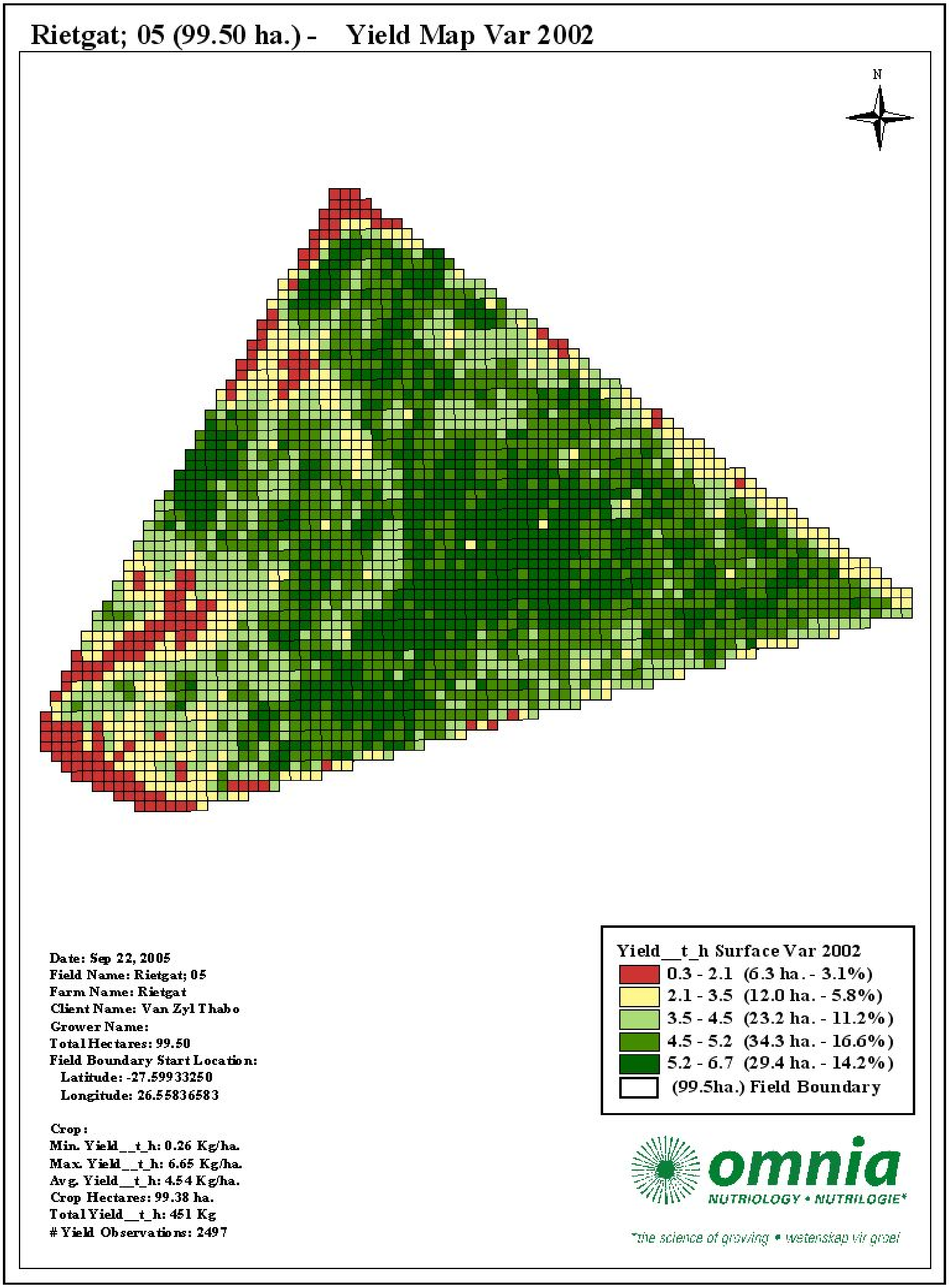

Figure 4.1b: VR Yield map 2002/2003

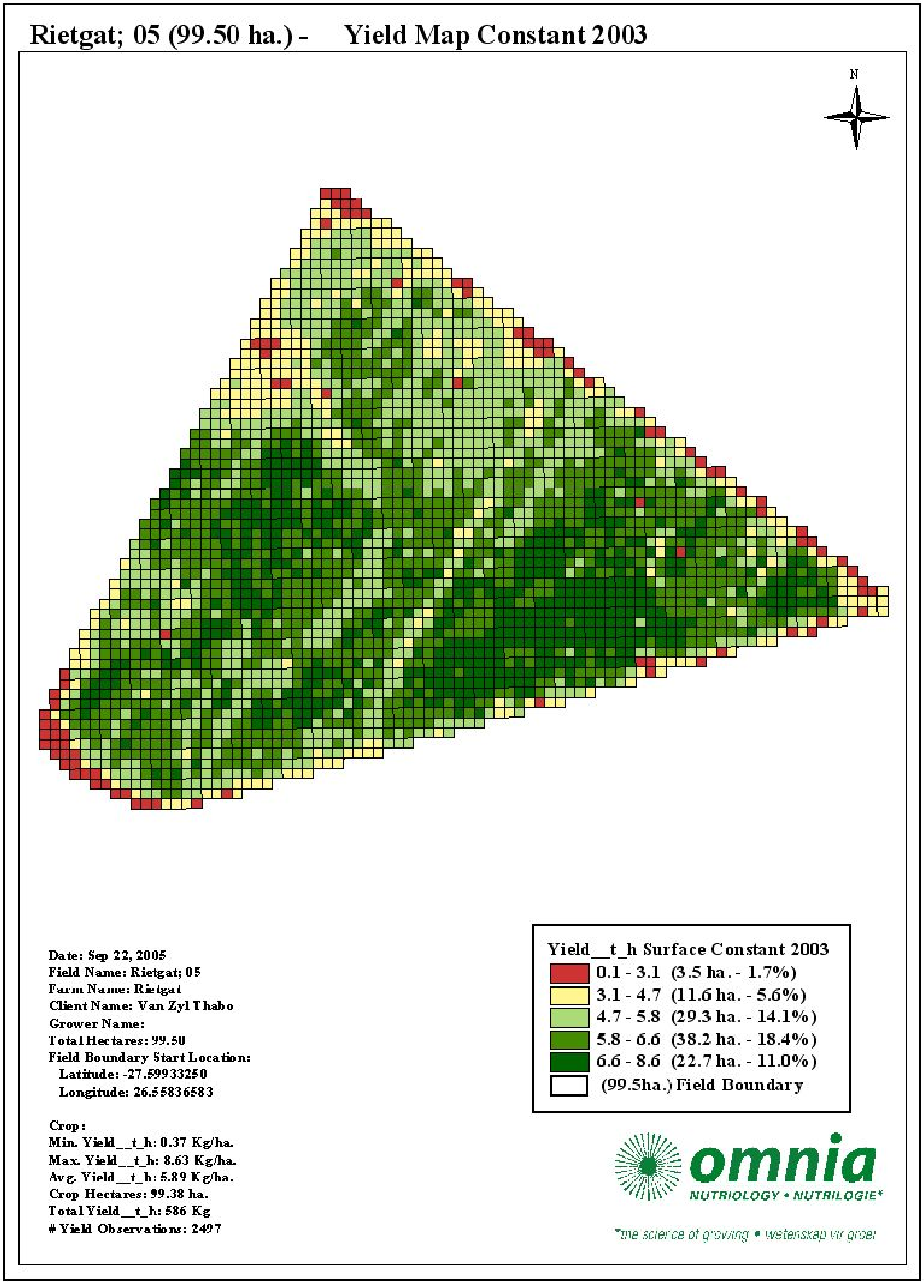

Figure 4.1c: SR Yield map 2003/2004

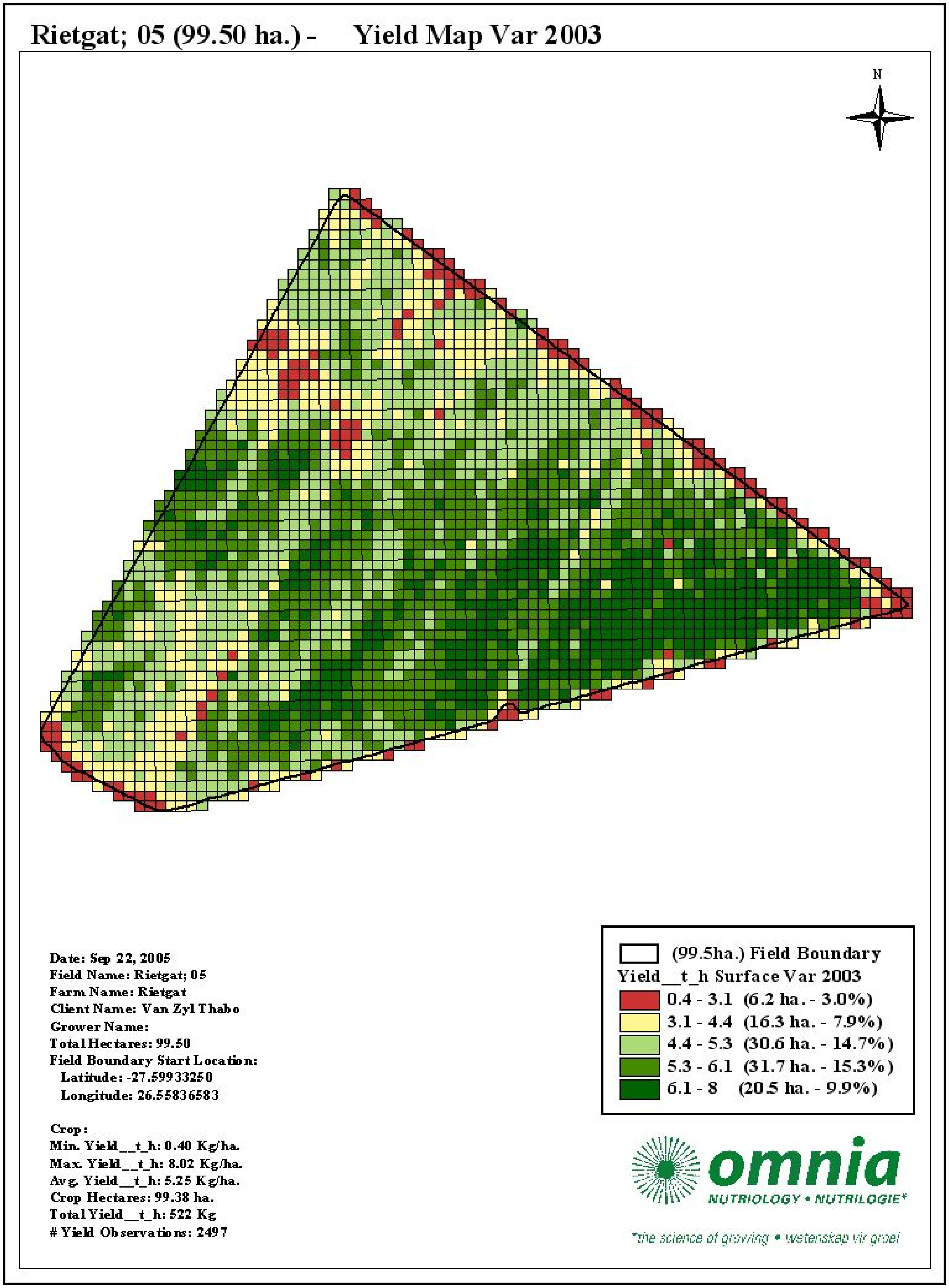

Figure 4.1d: VR Yield map 2003/2004

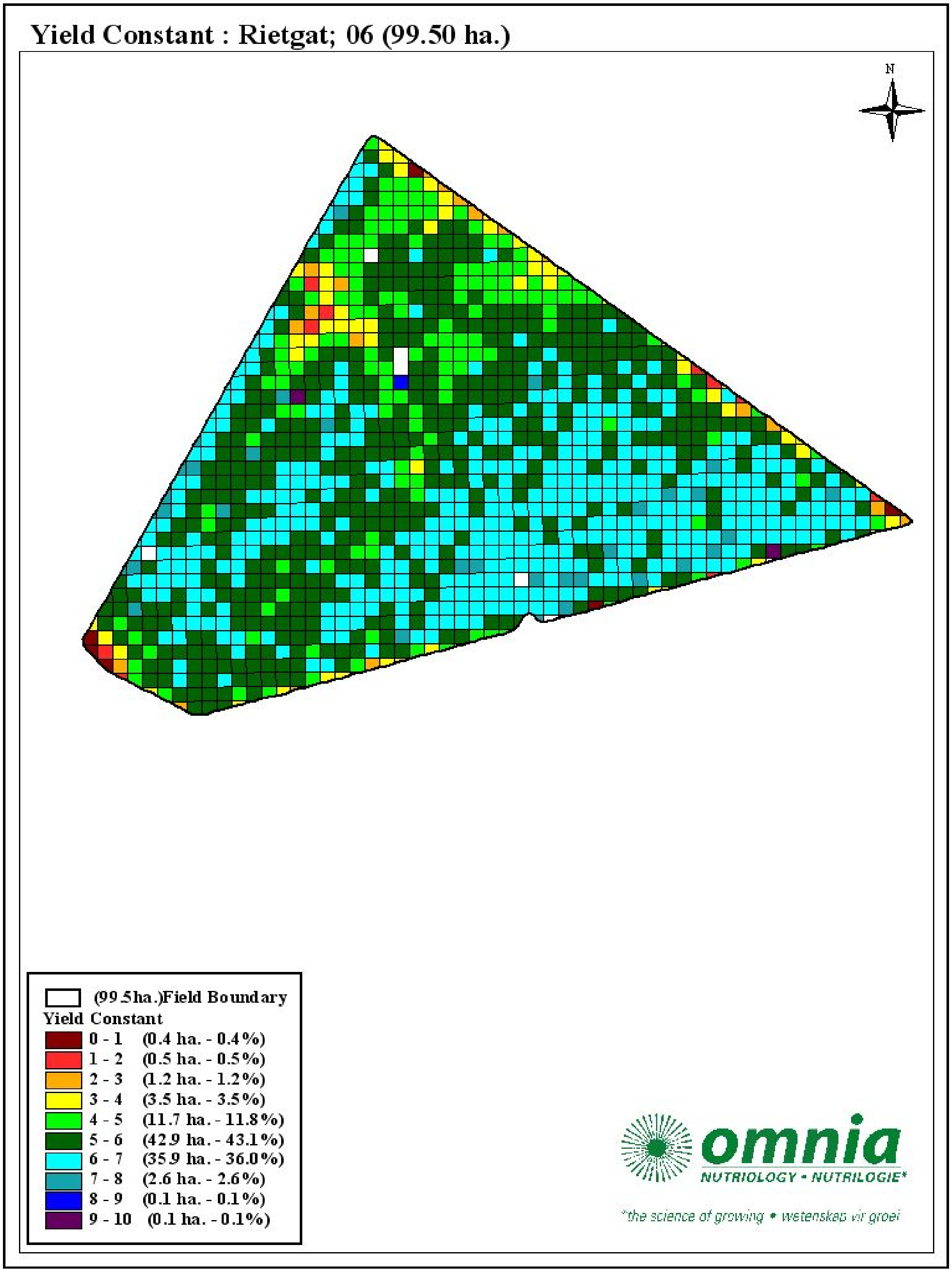

Figure 4.1e: SR Yield map 2004/2005

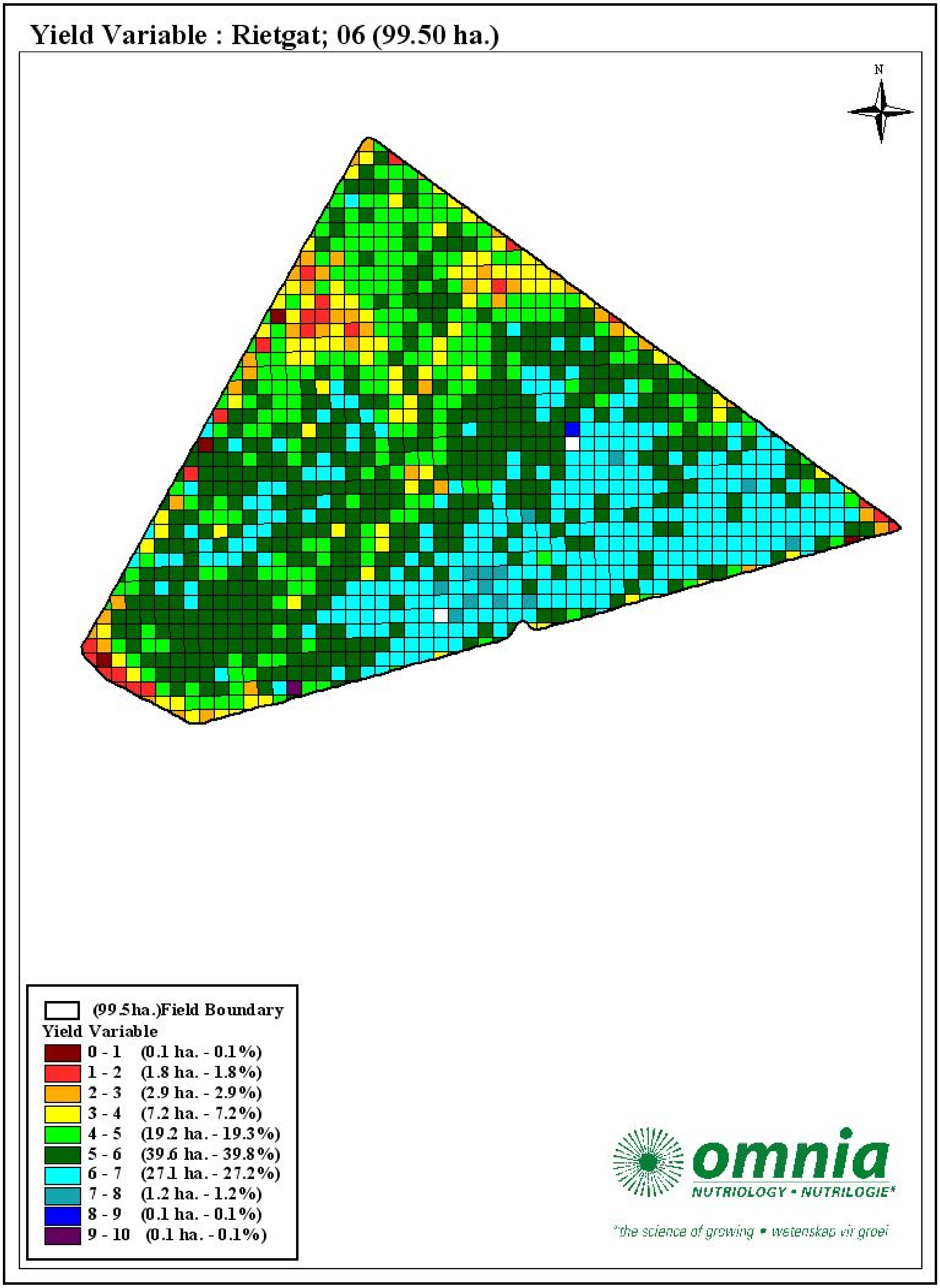

Figure 4.1f: VR Yield map 2004/2005

Annexure 4.2

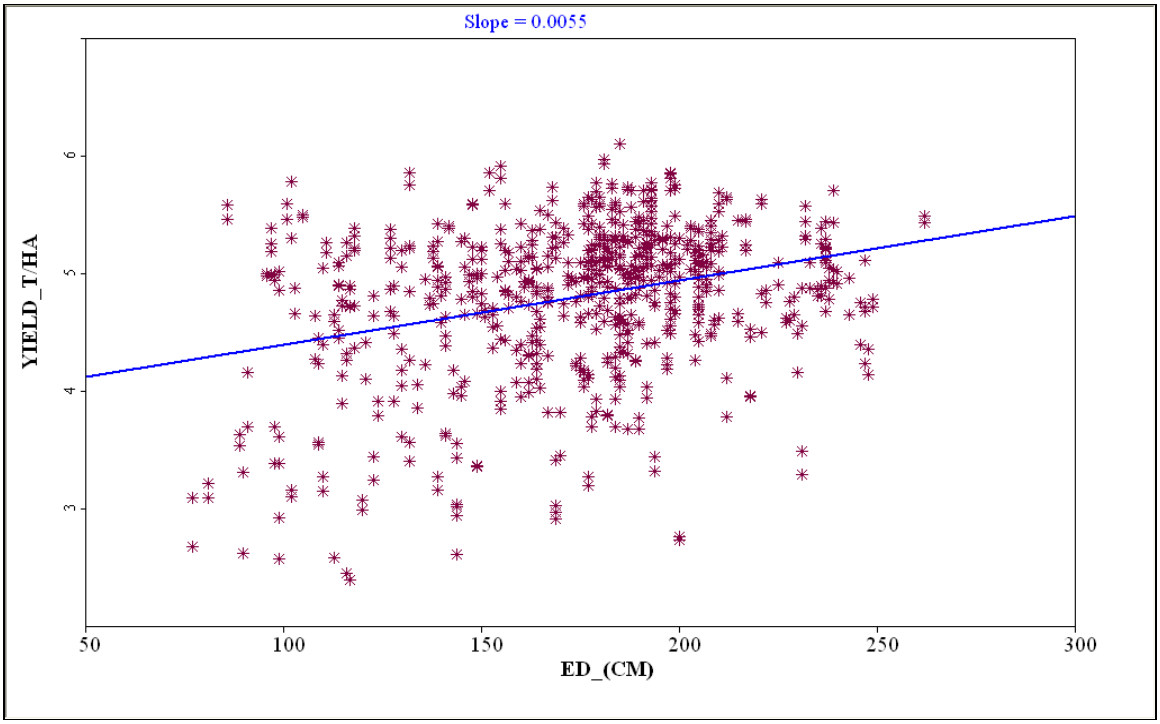

Figure 4.2a: Yield -Effective depth scatter plot – 2002/2003

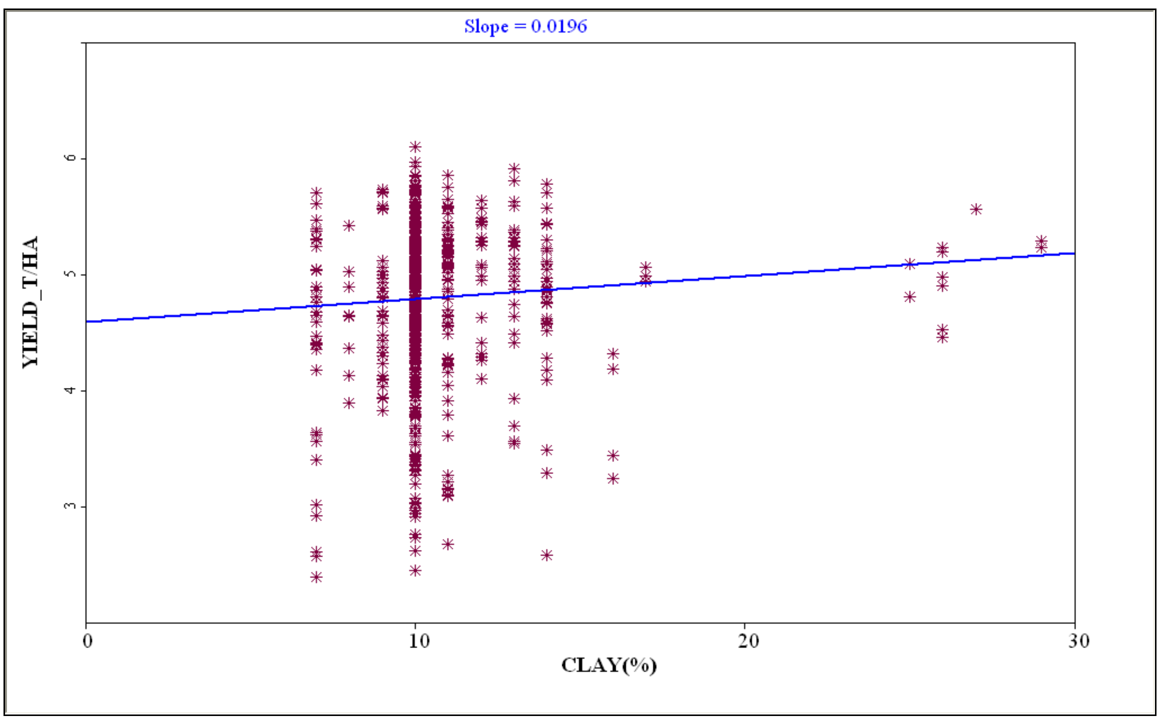

Figure 4.2b: Yield-Clay content scatter plot – 2002/2003

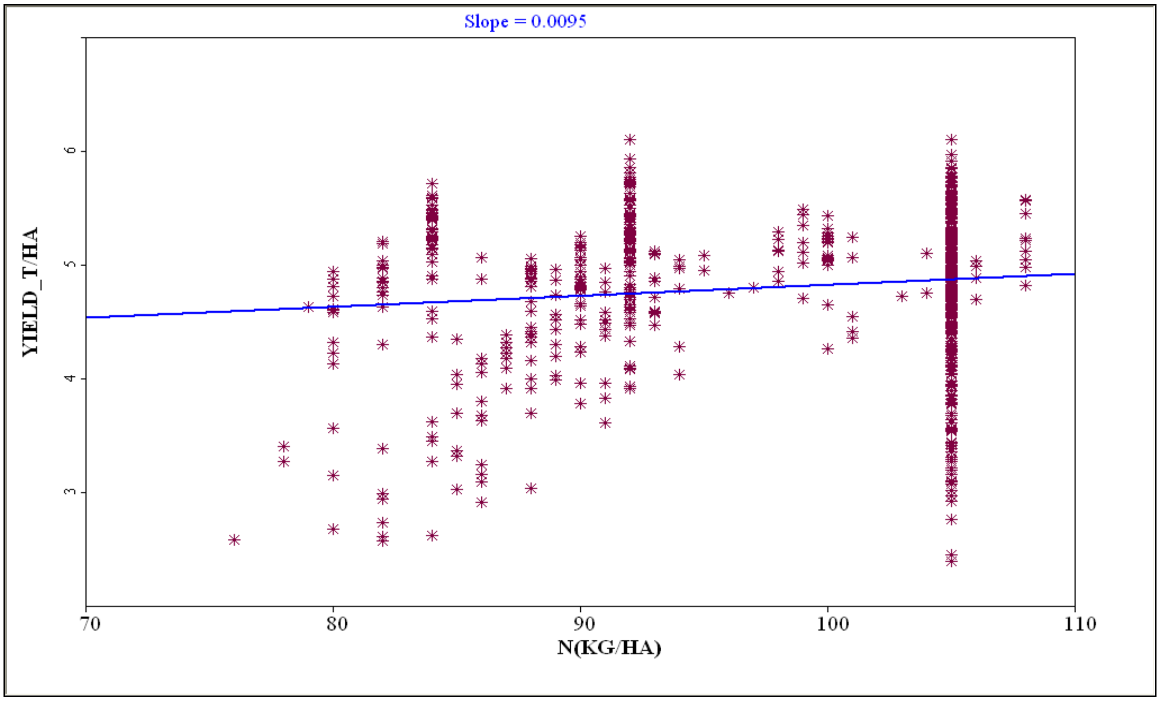

Figure 4.2c: Yield-Nitrogen scatter plot – 2002/2003

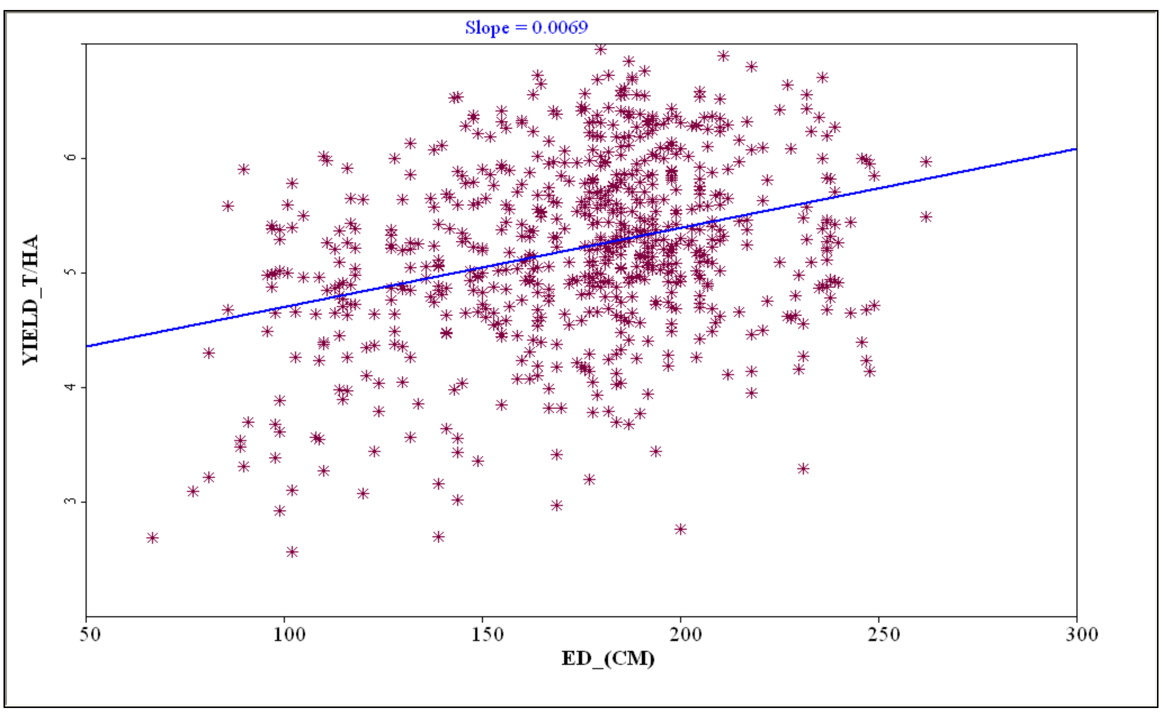

Figure 4.2d: Yield-Effective depth scatter plot – 2003/2004

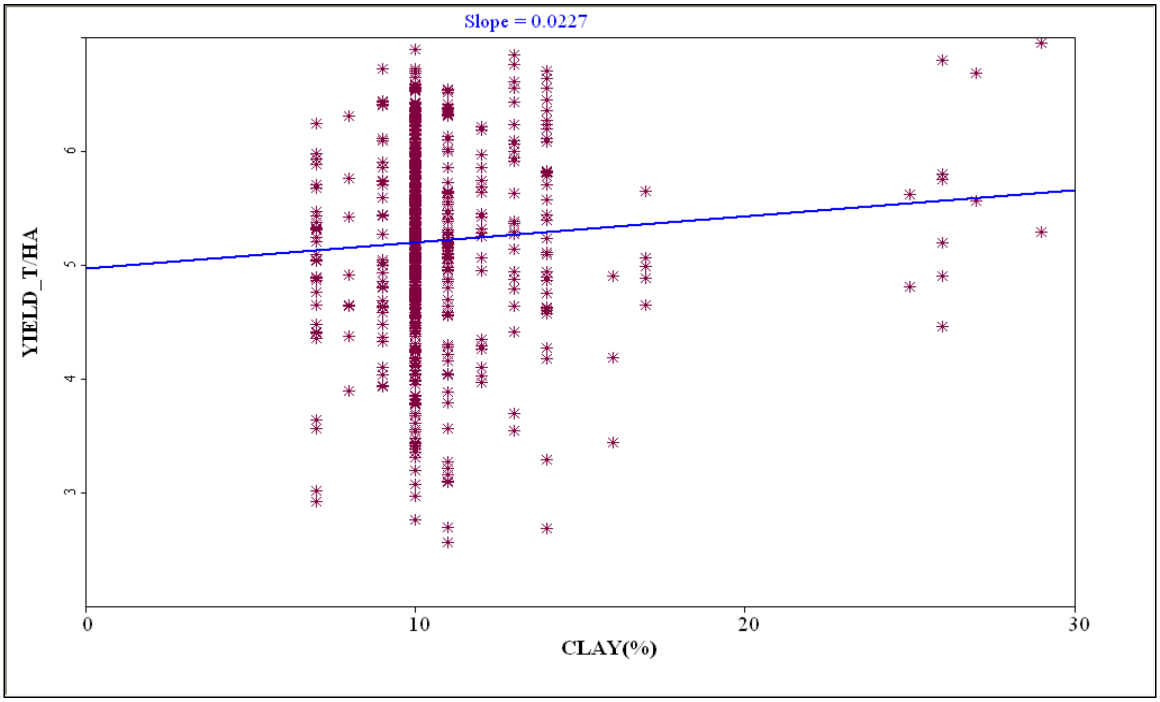

Figure 4.2e: Yield-Clay scatter plot – 2003/2004

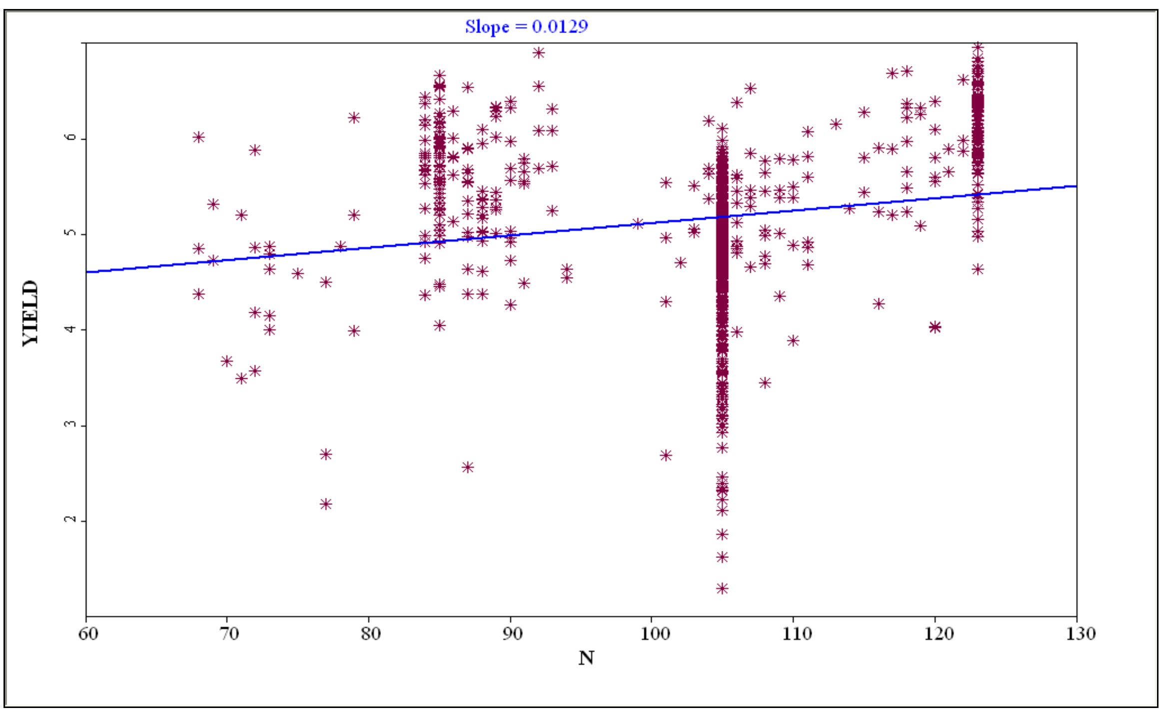

Figure 4.2f: Yield-Nitrogen scatter plot – 2003/2004

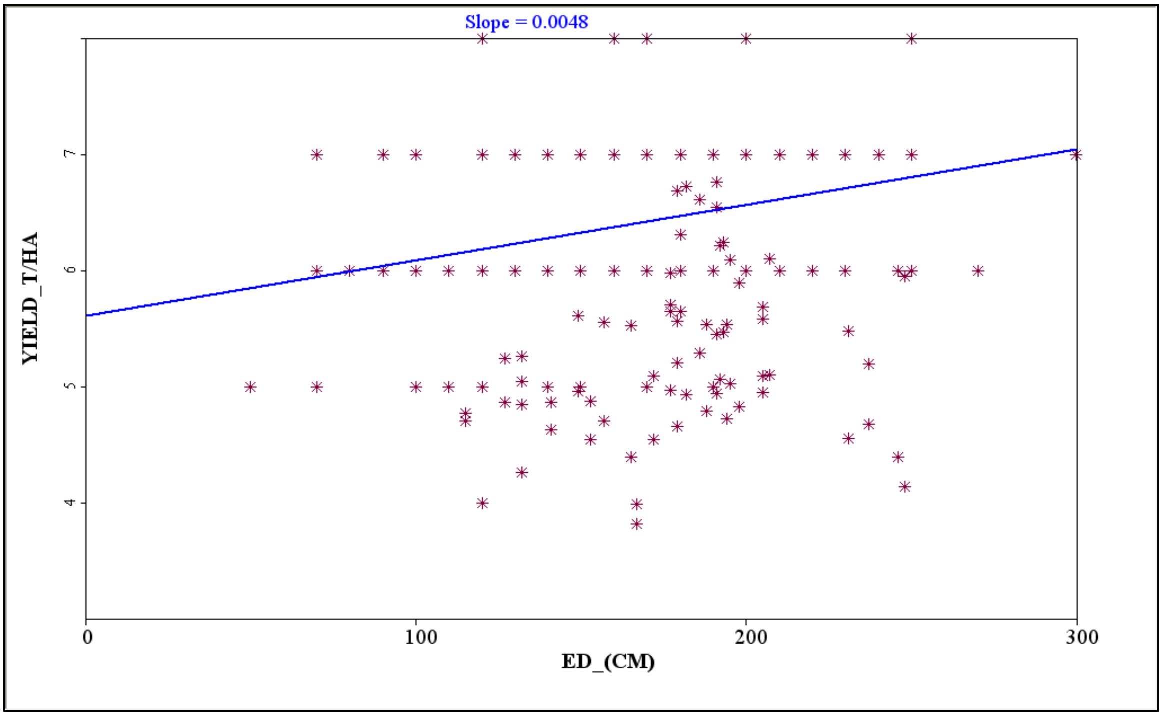

Figure 4.2g: Yield-Effective depth scatter plot – 2004/2005

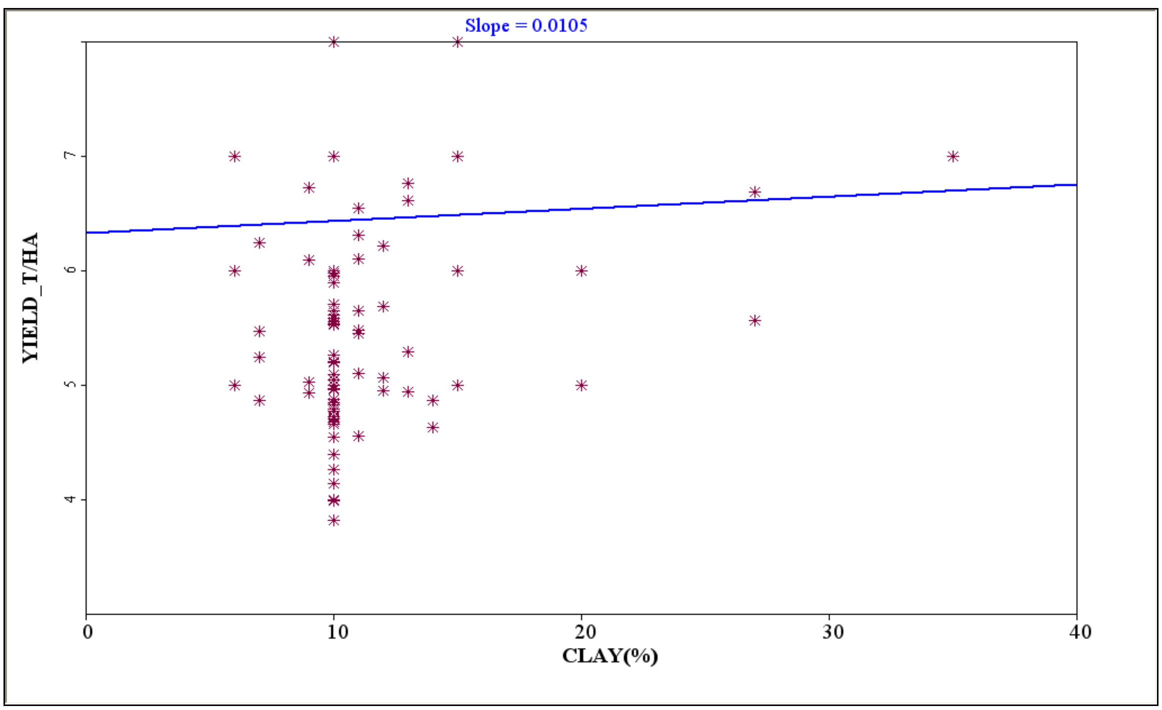

Figure 4.2h: Yield-Clay content scatter plot – 2004/2005

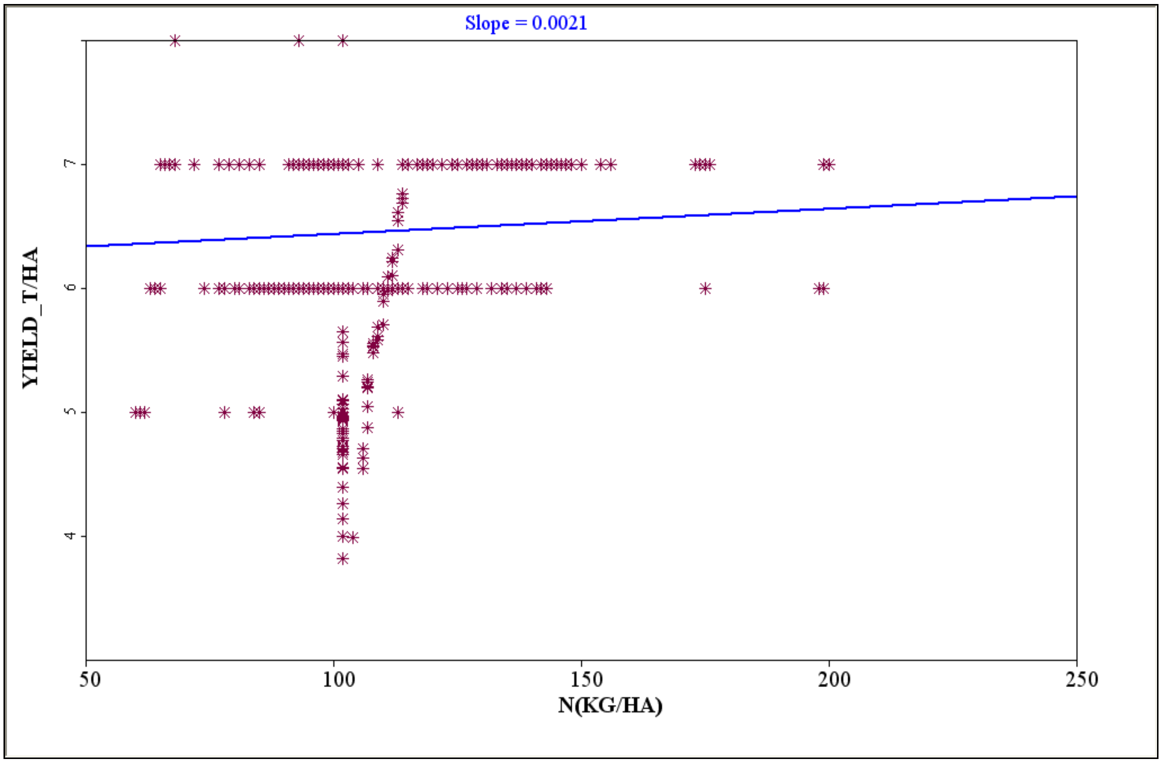

Figure 4.2i: Yield-Nitrogen scatter plot – 2004/2005

Annexure 4.3

ANNEXURE 4.3 (from Excel file)

Annexure 4.4

ANNEXURE 4.4 (Excel file)

Table 4.2:	Mean differences between clay contents of different zones
Table 4.3:	Mean differences between the effective soil depth of different zones
Table 4.4:	Mean yield differences per management zone between the three years

Annexure 5

TABLE 5.1: REGRESSION OUTPUT- BASELINE OLS MODEL YEAR 1 (2002/2003)

```
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : Year 1 map
Dependent Variable  :  YIELD_T/HA  Number of Observations:  723
Mean dependent var  :     4.80154  Number of Variables   :    8
S.D. dependent var  :    0.657489  Degrees of Freedom    :  715

R-squared           :    0.024401  F-statistic           :     2.55468
Adjusted R-squared  :    0.014849  Prob(F-statistic)     :   0.0132775
Sum squared residual:     304.921  Log likelihood        :    -713.789
Sigma-square        :    0.426463  Akaike info criterion :     1443.58
S.E. of regression  :    0.653041  Schwarz criterion     :     1480.25
Sigma-square ML     :    0.421744
S.E of regression ML:    0.649418
-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT      4.886875      0.1154425        42.3317    0.0000000
         TRT     0.1316435      0.1706513      0.7714181    0.4407118
          Z1    0.04416978      0.1403285      0.3147599    0.7530568
          Z2    -0.1556662      0.1251805      -1.243534    0.2140779
          Z3   -0.05749228      0.1363522     -0.4216453    0.6733965
      TRT_Z1    -0.1162409      0.2027273     -0.5733858    0.5665626
      TRT_Z2    -0.1745007      0.1838703     -0.9490422    0.3429148
      TRT_Z3    -0.1218157      0.1999955     -0.6090922    0.5426500
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   16.14656
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2           161.8017        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     7           81.25222        0.0000000
Koenker-Bassett test   7           50.93478        0.0000000
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 35            N/A            N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX: Weight_2.GWT  (row-standardized weights)
TEST                           MI/DF       VALUE          PROB
Moran's I (error)            0.883961    16.7921781      0.0000000
Lagrange Multiplier (lag)        1      325.8124527      0.0000000
Robust LM (lag)                  1       42.6400767      0.0000000
Lagrange Multiplier (error)      1      283.6484248      0.0000000
Robust LM (error)                1        0.4760487      0.4902175
Lagrange Multiplier (SARMA)      2      326.2885015      0.0000000
========================= END OF REPORT ==============================
```

TABLE 5.2: REGRESSION OUTPUT- BASELINE SER MODEL YEAR 1 (2002/2003)

```
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set             : Year 1 map
Spatial Weight       : Weight_2.GWT
Dependent Variable   :        YIELD  Number of Observations:  723
Mean dependent var   :     4.801535  Number of Variables   :    8
S.D. dependent var   :     0.657489  Degree of Freedom     :  715
Lag coeff. (Lambda) :     0.658735

R-squared            :     0.767409  R-squared (BUSE)      : -
Sq. Correlation      : -             Log likelihood        : -400.336098
Sigma-square         :     0.100547  Akaike info criterion :  816.672
S.E of regression    :     0.317092  Schwarz criterion     :  853.339470

-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
    CONSTANT      4.73639     0.07244958       65.37498    0.0000000
         TRT    0.1347469     0.05312675       2.536328    0.0112022
          Z1   -0.01636605    0.08555802      -0.191286    0.8483015
          Z2   -0.01611875    0.07491222     -0.2151685    0.8296360
          Z3   0.04425611     0.07952732      0.5564894    0.5778763
      TRT_Z1   -0.1085864     0.06308922      -1.721156    0.0852225
      TRT_Z2   -0.1817552     0.05713727      -3.181027    0.0014677
      TRT_Z3   -0.1353431     0.06234899      -2.170734    0.0299512
      LAMBDA    0.6587349      0.0149038       44.19912    0.0000000
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                        DF      VALUE         PROB
Breusch-Pagan test                          7        40.97292     0.0000008

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX: Weight_2.GWT
TEST                                         DF     VALUE         PROB
Likelihood Ratio Test                        1      626.9058      0.0000000
========================= END OF REPORT ==============================
```

TABLE 5.3: REGRESSION OUTPUT- BASELINE OLS MODEL YEAR 2 (2003/2004)

SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

```
Data set            : Year 2 map
Dependent Variable  :  YIELD_T/HA  Number of Observations:  719
Mean dependent var  :     5.21145  Number of Variables   :    8
S.D. dependent var  :    0.793561  Degrees of Freedom    :  711

R-squared           :    0.274721  F-statistic           :  38.4731
Adjusted R-squared  :    0.267580  Prob(F-statistic)     : 7.4546e-046
Sum squared residual:     328.393  Log likelihood        :  -738.495
Sigma-square        :    0.461875  Akaike info criterion :   1492.99
S.E. of regression  :    0.679614  Schwarz criterion     :   1529.61
Sigma-square ML     :    0.456736
S.E of regression ML:    0.675823
```

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	4.877879	0.1183056	41.23116	0.0000000
TRT	0.7863636	0.1673094	4.700056	0.0000031
Z1	0.08255599	0.1438404	0.5739416	0.5661976
Z2	-0.1267056	0.1287499	-0.9841221	0.3253922
Z3	-0.06167626	0.1408643	-0.4378417	0.6616475
TRT_Z1	-0.08799245	0.2039116	-0.4315225	0.6662129
TRT_Z2	-0.07660275	0.181963	-0.4209799	0.6738915
TRT_Z3	0.279252	0.199593	1.399108	0.1622152

REGRESSION DIAGNOSTICS

MULTICOLLINEARITY CONDITION NUMBER 15.55384

TEST ON NORMALITY OF ERRORS

TEST	DF	VALUE	PROB
Jarque-Bera	2	127.8075	0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	7	47.69032	0.0000000
Koenker-Bassett test	7	30.97616	0.0000628

SPECIFICATION ROBUST TEST

TEST	DF	VALUE	PROB
White	35	N/A	N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE

FOR WEIGHT MATRIX: **Weight_800.GWT** (row-standardized weights)

TEST	**MI/DF**	**VALUE**	**PROB**
Moran's I (error)	0.476088	35.7271362	0.0000000
Lagrange Multiplier (lag)	1	1206.3716802	0.0000000
Robust LM (lag)	1	8.7128825	0.0031597
Lagrange Multiplier (error)	1	1213.2301714	0.0000000
Robust LM (error)	1	15.5713737	0.0000794
Lagrange Multiplier (SARMA)	2	1221.9430539	0.0000000

========================= END OF REPORT ==============================

TABLE 5.4: REGRESSION OUTPUT- BASELINE SER MODEL YEAR 2 (2003/2004)

```
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : Year 2 maps
Spatial Weight      : Weight_800.GWT
Dependent Variable  :       YIELD  Number of Observations:  721
Mean dependent var  :    5.212150  Number of Variables   :    8
S.D. dependent var  :    0.792591  Degree of Freedom     :  713
Lag coeff. (Lambda) :    0.856322

R-squared           :    0.625545  R-squared (BUSE)      : -
Sq. Correlation     : -            Log likelihood        : -532.050352
Sigma-square        :    0.235233  Akaike info criterion : 1080.1
S.E of regression   :    0.485008  Schwarz criterion     : 1116.745817

----------------------------------------------------------------------
    Variable    Coefficient     Std. Error     z-value      Probability
----------------------------------------------------------------------
    CONSTANT      4.84456      0.1538919       31.48028    0.0000000
         TRT    0.7856226      0.1130909       6.946823    0.0000000
          Z1   -0.01327598     0.1137816     -0.1166795    0.9071139
          Z2   -0.05308311     0.09936967     -0.5341983   0.5932043
          Z3   -0.08634867     0.1107201     -0.7798822    0.4354601
      TRT_Z1   -0.06755906      0.137672     -0.4907248    0.6236211
      TRT_Z2   -0.05788107     0.1229554     -0.4707485    0.6378203
      TRT_Z3    0.2683828      0.1349506       1.988749    0.0467288
      LAMBDA     0.856322     0.03056131       28.01981    0.0000000
----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                     DF      VALUE          PROB
Breusch-Pagan test                       7        42.91611     0.0000003

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX: Weight_800.GWT
TEST                                     DF      VALUE          PROB
Likelihood Ratio Test                    1        415.0341     0.0000000
========================= END OF REPORT ==============================
```

TABLE 5.5: REGRESSION OUTPUT- BASELINE OLS MODEL YEAR 3 (2004/2005)

```
SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION
Data set            : Year 3 map
Dependent Variable  :  YIELD_T/HA  Number of Observations:  728
Mean dependent var  :     6.43941  Number of Variables   :    8
S.D. dependent var  :    0.712049  Degrees of Freedom    :  720

R-squared           :    0.358637  F-statistic           :     57.5155
Adjusted R-squared  :    0.352401  Prob(F-statistic)     :           0
Sum squared residual:     236.731  Log likelihood        :    -624.078
Sigma-square        :    0.328793  Akaike info criterion :     1264.16
S.E. of regression  :    0.573405  Schwarz criterion     :     1300.88
Sigma-square ML     :     0.32518
S.E of regression ML:    0.570246
-----------------------------------------------------------------------
    Variable    Coefficient     Std. Error    t-Statistic   Probability
-----------------------------------------------------------------------
    CONSTANT      4.877879     0.09981699       48.86822    0.0000000
         TRT     0.7863636      0.1411625       5.570626    0.0000000
          Z1      1.663788      0.1205405       13.80273    0.0000000
          Z2       1.53355      0.1088221       14.09226    0.0000000
          Z3      1.995539      0.1188502       16.79037    0.0000000
      TRT_Z1    -0.7505655      0.1706585      -4.398055    0.0000126
      TRT_Z2    -0.7784374      0.1535363      -5.070055    0.0000005
      TRT_Z3    -0.7863636      0.1680796       -4.67852    0.0000035
-----------------------------------------------------------------------


REGRESSION DIAGNOSTICS
MULTICOLLINEARITY CONDITION NUMBER   15.74793
TEST ON NORMALITY OF ERRORS
TEST                  DF          VALUE            PROB
Jarque-Bera            2           42.69568        0.0000000

DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                  DF          VALUE            PROB
Breusch-Pagan test     7           33.45225        0.0000218
Koenker-Bassett test   7           28.17352        0.0002045
SPECIFICATION ROBUST TEST
TEST                  DF          VALUE            PROB
White                 35            N/A            N/A

DIAGNOSTICS FOR SPATIAL DEPENDENCE
FOR WEIGHT MATRIX: Weight_1.GWT  (row-standardized weights)
TEST                           MI/DF       VALUE          PROB
Moran's I (error)            0.665596    36.0124017      0.0000000
Lagrange Multiplier (lag)        1      883.8913501      0.0000000
Robust LM (lag)                  1        0.0767509      0.7817498
Lagrange Multiplier (error)      1      1266.8992264     0.0000000
Robust LM (error)                1      383.0846272      0.0000000
Lagrange Multiplier (SARMA)      2      1266.9759772     0.0000000
========================= END OF REPORT ==============================
```

TABLE 5.6: REGRESSION OUTPUT- BASELINE SER MODEL YEAR 3 (2004/2005)

```
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set             : Year 3 map
Spatial Weight       : Weight_1.GWT
Dependent Variable   :  YIELD_T/HA  Number of Observations:  728
Mean dependent var   :    6.439409  Number of Variables   :    8
S.D. dependent var   :    0.712049  Degree of Freedom     :  720
Lag coeff. (Lambda) :    0.915684

R-squared            :    0.821928  R-squared (BUSE)      : -
Sq. Correlation      : -            Log likelihood        : -225.065833
Sigma-square         :    0.090285  Akaike info criterion :     466.132
S.E of regression    :    0.300474  Schwarz criterion     :  502.854074

-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
    CONSTANT     4.661919      0.1447953       32.19662    0.0000000
         TRT    0.7875385       0.066724       11.80293    0.0000000
          Z1     1.810325     0.07550058       23.97763    0.0000000
          Z2     1.830309     0.06593084       27.76104    0.0000000
          Z3     1.900605     0.07306426       26.01279    0.0000000
      TRT_Z1   -0.7778204     0.08064283      -9.645253    0.0000000
      TRT_Z2   -0.7926782     0.07263191      -10.91364    0.0000000
      TRT_Z3   -0.7875385     0.07925728      -9.936481    0.0000000
      LAMBDA    0.9156842     0.01663601       55.04231    0.0000000
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                    DF      VALUE          PROB
Breusch-Pagan test                      7        31.95059     0.0000415

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX: Weight_1.GWT
TEST                                    DF      VALUE          PROB
Likelihood Ratio Test                   1        798.0251     0.0000000
========================= END OF REPORT ==============================
```

Annexure 6

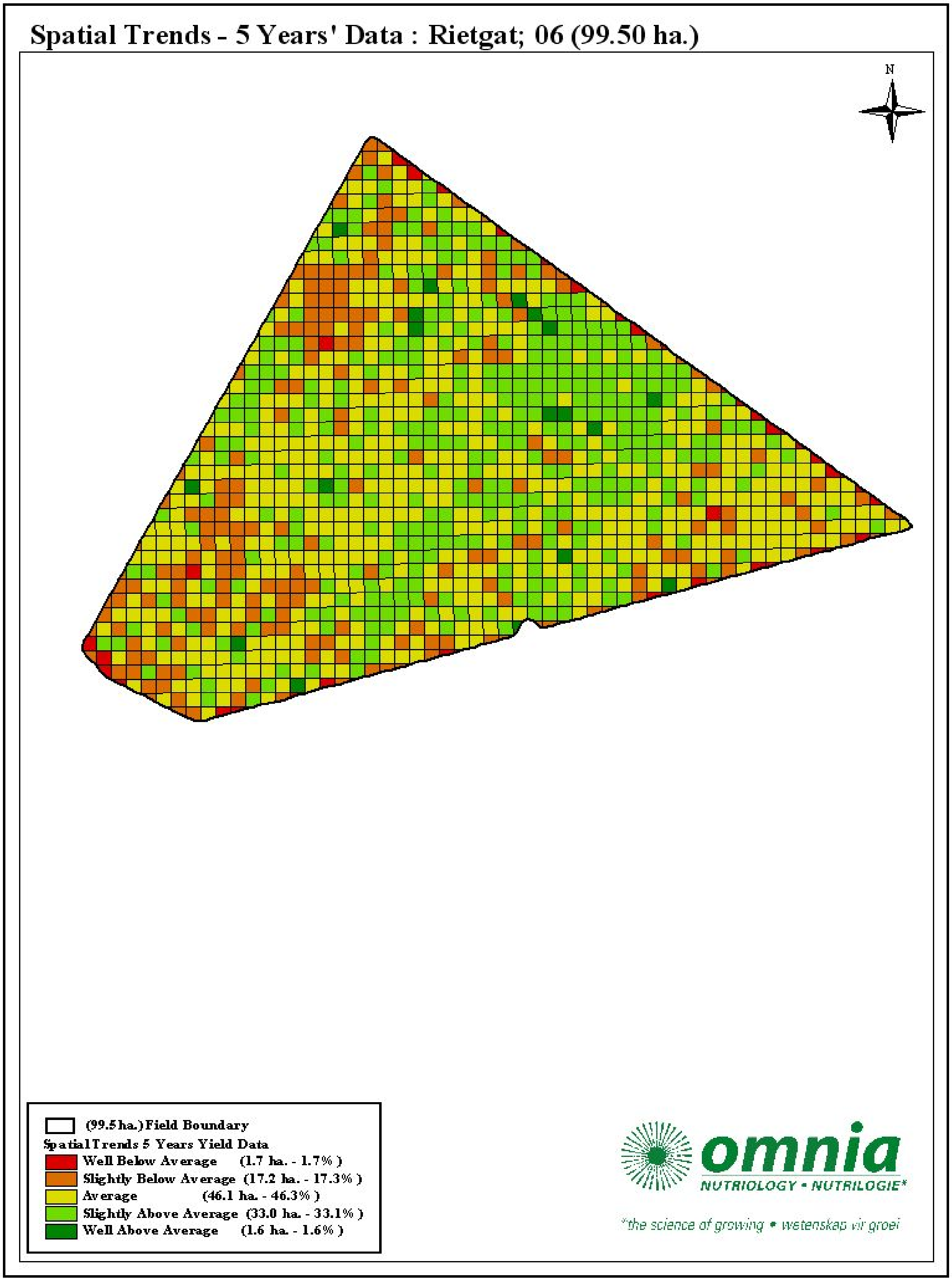

Figure 6.1: Spatial trends- 5 year data

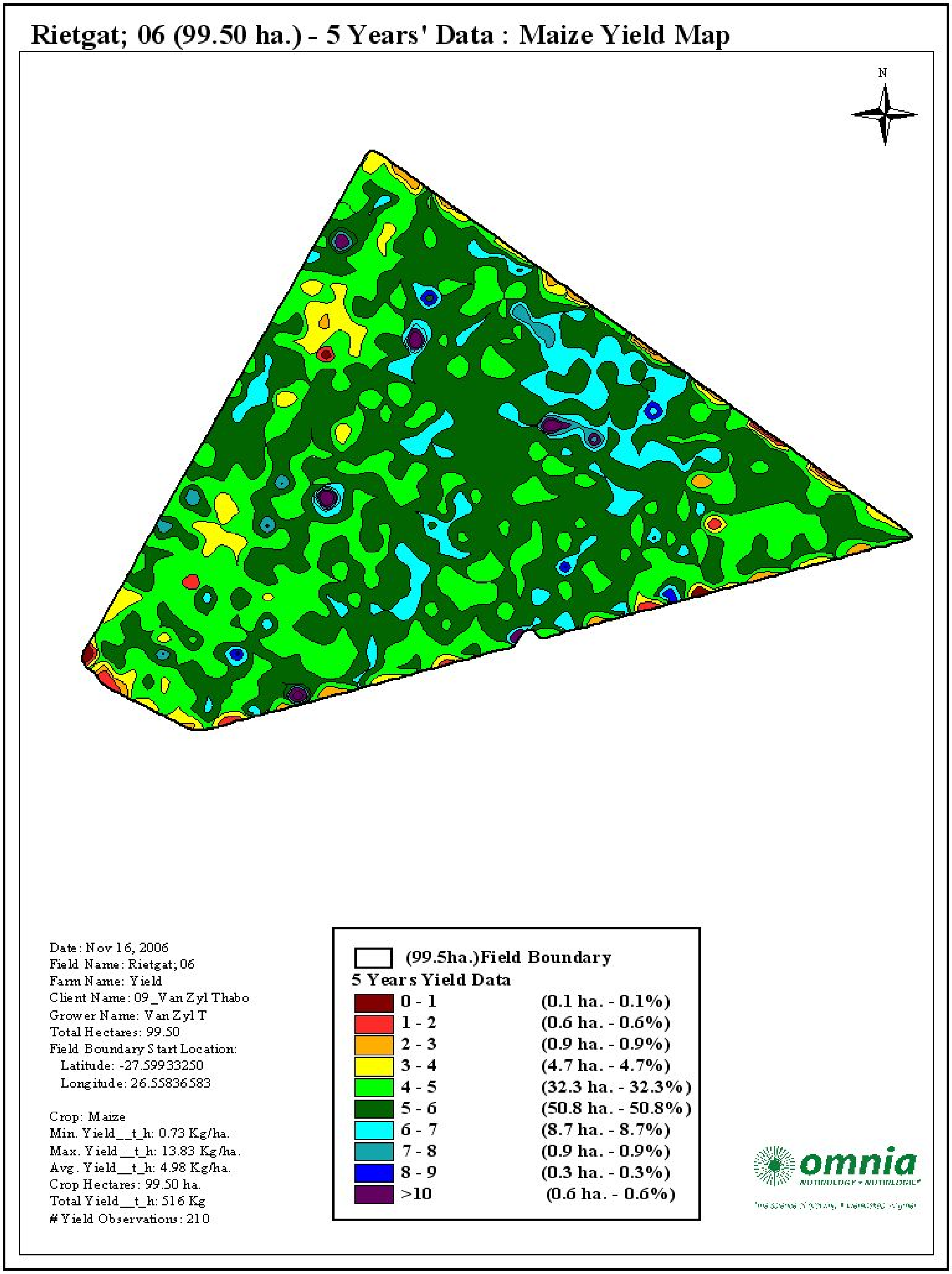

Figure 6.2: 5-year data yield map (2001-2005)